Multimodalities in Metadata: Gaia Gate

The execution and analysis of external contextual metadata in data-driven renderings and multimodal media

by TYSON MOLL

A thesis presented to OCAD University in partial fulfillment of the requirements for the degree of Master of Design in Digital Futures

Toronto, ON, Canada, 2020

Abstract

Metadata is information about objects. Existing metadata standards seldom describe details concerning an object's context within an environment; this thesis proposes a new concept, **external contextual metadata (ECM)**, examining metadata, digital photography, and mobile interface theory as context for a proposed multimodal framework of media that expresses the internal and external qualities of the digital object and how they might be employed in various use cases. The framework is binded to a digital image as a singular object. Information contained in these 'images' can then be processed by a renderer application to reinterpret the context that the image was captured, including non-visually. Two prototypes are developed through the process of designing a renderer for the new multimodal data framework: a proof-of-concept application and a demonstration of 'figurative' execution (titled 'Gaia Gate'), followed by a critical design analysis of the resulting products.

Land Acknowledgement

I acknowledge that this thesis was written on the traditional territory of multiple nations including the Mississaugas of the Credit, the Anishinaabe, the Chippewa, the Haudenosaunee, the Wendat peoples in addition to the many peoples (First Nations, Inuit, Métis) who call this land home today. I also acknowledge that Toronto is covered by Treaty 13 with the Mississaugas of the Credit.

Dedication and Thanks

To my late grandfather, Peter Moll: an avid collector and documenter of stamps and cameras, and an all-round rosy, well-humoured gentleman.

I would also like to thank my Digital Futures cohort, the SAD GONG Design Collective community, my advisors Cindy Poremba and Alexis Morris, Isabelle and the Haines family, as well as my parents and friends for their support and kindness throughout the process of preparing this document.

Table of Contents

List of Figures

All figures captured or created by the author unless otherwise noted / cited on their respective pages.

Glossary of Terms

This thesis uses several new and esoteric terms as part of its discourse. Some of these terms are listed below with their most relevant section highlighted in brackets.

contextual metadata framework: [Section 4.0-4.2] the combination of both existing, 'internal' contextual metadata and the newly proposed external contextual metadata as a new metadata framework.

external contextual metadata (ECM): [Section 4.1] metadata concerning the *external* context of an object. In other words, information concerning details related to where an object is situated in space and time external to the object itself.

Exif (Exchangeable image file format): [Section 2.3.1] a common format of metadata associated with digital photography, which includes fields for capture device [camera] settings, time and day, GPS location, and more.

metadata: [Section 2.1] information about objects, generally in a digital context.

multimodal: possessing multiple modalities (modes of experience)

multimodal frame: a photographic image with additional, non-visual modes of experience (alluding to a slice of a film).

renderer: [Section 1.0, 4.0] a software application that represents data in a sensory capacity (not necessarily limited to visualizations)

rendering application: see *renderer*

steganography: [Section 4.2.1] the encoding of hidden information into digital objects in a manner that cannot be read through conventional use of the digital object.

1.0 Introduction

A large quantity of circumstantial information surrounds a photograph. Where was it taken? What was the weather like? The sound of seagulls and the smell of the ocean breeze are absent from the frame, however evocative their representation may be. But a photograph is but a plane with a framed image. We are free to physically manipulate the print, inscribe information about it, observe it under varying lighting conditions and translate it through space and time. A digital photograph is very much the same. We have access to information in a digital image file's *metadata* describing the image or details about the device that captured the photograph. However, existing metadata standards for digital images are very limited in their description of the external context concerning an object. Metadata standards enforce a limited field of description compared to the seemingly infinite possibilities extended to us with physically modifiable materials. **I seek to propose a new framework for digital media metadata that integrates a description of a digital media object's context. I call this *external contextual metadata* (ECM)[1], metadata that describes information about the context of an object that are external to the object's inherent properties.**

[1] Elaborated upon in Chapter 4: Theoretical Digital Framework

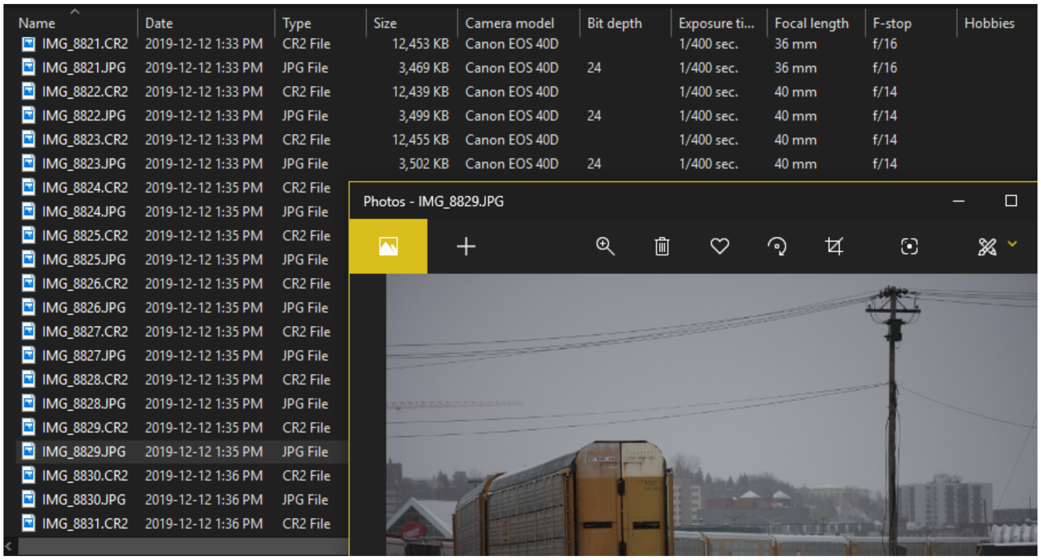


Figure 1: Digital photographs captured with a DSLR camera, listed with several metadata tags (screenshot from explorer.exe, Windows 10)

Our digital images can tell us about their global position, time, and their capture device settings; information asserted to be useful to the average photographer. This data is largely restricted to information within the visual frame or information describing the photo as an object, requiring that supplementary information be ascribed to meta-tags by users and artificial intelligence. What interests me (and explicitly serves as my research question) is how metadata describing the context external to a digital object might supplement existing digital metadata to further elaborate on the context surrounding a digital artefact as presented through a rendering software.

I will specifically be focusing on metadata associated with digital images with locative aspects (including digital photography), but it may be presumed that the aspects that I describe and investigate concerning the relationship between metadata and context can be applied beyond the digital image and locative properties. To this end, the photograph is no longer merely an image but the combination of multisensory and non-sensory modalities, prompting the need for a new set of terminology to describe this digital content: the 'multimodal frame' (the digital object in central to this inquiry) and its rendering software, the 'renderer' (usurping ocular centric photo gallery applications).

As part of this research investigation, I will present a developed prototype, dubbed 'Gaia Gate', demonstrating multimodal frames through the processes of capture and rendering, complemented with insights acquired through a Research-Creation process. I will also discuss how the multimodal framework might integrate with existing society as modelled by several user groups and their respective use cases. This method will establish a means of evaluating and analyzing the proposal and prototype, as more technically-nuanced[2] means of evaluating the prototype are beyond scope and are challenging to directly investigate due to the absence of this proposed metadata

[2] e.g. Statistical investigation of new media usage, wide-scale testing of the metadata format, a thoroughly developed standard for metadata.

standard in modern society; the focus of this thesis is directed towards the advantages for an application developer as is appropriate for a design thesis.

Ultimately, this thesis serves to provoke a discussion on the capture, encapsulation, and display of secondary, external contextual metadata (ECM) within a single digital file as the basis for a multimodal contextual metadata framework for 'multimodal frames', contributing to a dialogue towards the next major advancement in metadata standards. Up until this point, there is little evidence of a publicly-available metadata format that acknowledges ECM in this manner beyond the umbrella term of 'arbitrary data', data to be custom-indexed or tagged by users of extensible metadata formats like Adobe's XMP without a standardized format. The contributions of the work are directed towards photography, data visualization, application development, and new media studies, although the expected ubiquitous nature of a change to metadata standards could affect a broader scope of use cases including digital histories, software engineering, and surveillance technologies.

1.1 Chapter Overview

The chapters of this thesis are outlined in brief below:

- Chapter 1 (Introduction) briefly introduces the topics of this thesis, the research questions concerning the idea of external context as metadata, as well as the concept of a 'renderer'.
- Chapter 2 (Literature and Context) provides a background review for the investigation covering the topics of metadata, data visualization, digital photography, and mobile interface theory.
- Chapter 3 (Methodologies) covers the research methodologies that will largely be explored in the Creative Process sections, namely Research-Creation and Discursive Design. Research-Creation reflects the feedback loop relationship between the research and prototyping process whereas Discursive Design is employed through the analysis of several user groups and their respective use cases for multimodal frames, facilitating the study and discussion of external contextual metadata were it to be ubiquitous within present society.
- Chapter 4 (Theoretical Framework) builds upon the foundations covered in the context and literature review and establishes a niche in metadata standards ripe for investigation: external contextual metadata (ECM). The concept of a multimodal frame is introduced, including a framework for contextual metadata scoped to accommodate the research, as well as several design considerations from both a technical standpoint and regarding related practices.

- Chapter 5 (Use Case Discussion) outlines the relationships between the prototypes and the Discursive Design investigation and establishes the worldbuilding context in which the user groups will be discussed. The Application Developer, The Photographer, and The End-User covers the three user groups directly connected to multimodal frames and their renderers, whereas The Archivist, Surveillance, and Automata represent a selection of topics concerning how multimodal frames might be indirectly used as well as how existing and future technological practices may take advantage of the new metadata format.
- Chapter 6 (Renderer Prototyping) embodies the Application Developer use case in the prototyping process and includes several insights related to their practice. This chapter features a Proof-of-Concept prototype that demonstrates the multimodal capture process and a Demonstrative Prototype ('Gaia Gate') examining one possible application of rendering multimodal frames as a high-fidelity prototype. The chapter concludes with an evaluation of the prototype based on criteria from Chapter 4.
- Chapter 7 (Conclusion) wraps up the thesis with a discussion of the creative process in its entirety, concluding remarks, as well as several avenues for future investigation.

2.0 Review of Literature & Context

The goal of this literature review is to establish a background for the investigation of External Contextual Metadata (ECM) and its use in digital rendering applications, primarily for use with mobile devices. The actual definition for contextual metadata and proposals for its framework will be established in subsequent chapters. For the sake of clarity and brevity, attention will be directed to the most relevant elements of these studies. For those interested in a list and comparison of media discussed in this thesis directly related to the prototype such as games, applications, artworks, and locative activities, please refer to Section 6.4.

The topics that will be covered are as follows:

1. *Metadata (and Data)*: information about objects, or data about data, with a review the history of metadata and its standardized usage.
2. *Data Visualization (and Data Analytics)*: the process of creating visual expressions of data, including the benefits of the practice for cognition and its association with general Data Analytics.
3. *Digital Photography (and the Digital Image)*: photography for digital devices, digital image formats, and a brief note on metadata retention for images on social media.

4. *Mobile Interfaces (and Hybrid Spaces)*: the mobile device as an interface between real and virtual spaces, and some implications of this technology.

2.1 Metadata and Data

Data is information, by definition either known or assumed to be factual (Oxford n.d.). In the 1640s, *datum* became understood as “a fact given or granted” generally in reference to mathematical or numerical calculations, and its association with computer technology emerged in 1946 (Harper n.d.). Similarly, *metadata* is information, or data, that describes another object (Pomerantz 2015, 26). In 1995, the Dublin Core was established and defined a standardized set of 15 ‘core’ metadata elements[3] to describe all manners of ‘networked resources’ (digital objects shared across networks) (Pomerantz 2015, 67). The Dublin Core was one of the most important initiatives for networked metadata as it established a baseline standard for metadata descriptions that became widely adopted and ubiquitous on the web (Pomerantz 2015, 71) and has since been succeeded by multiple modern metadata structures to this day.

---

[3] The elements of the Dublin Core are Contributor, Coverage, Creator, Date, Description, Format, Identifier, Language, Publisher, Relation, Rights, Source, Subject, Title, and Type. They also created a ‘qualified’ extension set of 40 Terms in addition to the 15 in the core set, which in general terms describe an object’s relationships and interactions with other objects (e.g. hasPart, isPartOf, modified, audience). (Pomerantz 2015, 81)

Metadata is both a property and subset of data, and its definitions tend to overlap with definitions of data. Colin Ware describes three categories of data:

- *Entities* (or, objects of interest)
- *Relations* (e.g. parent-of, part-of, etc.)
- and *Operations* (e.g. transformations, addition)

(Ware 2013, 25–26)

All three can have attributes (categorical, numerical, etc.) that describe their features. The domain of metadata includes these features as well as the *relations*, a category of data that effectively describes information about a connection between object(s), or simply data about data. For clarity's sake, attributes and relations will be considered *metadata*, entities will be considered *data*, and operations will be referred to independently of data as Ware himself expresses some uncertainty in its positioning as a form of data (2013). It is not an object's properties themselves, but rather a description of those properties that defines metadata; the indication that an apple is red rather than the redness of the apple itself. This draws from the origins of the term 'metadata' in Aristotle's metaphysics, which literally translates to 'after physics [or, after the study of nature]'; the study of the existence of things that have already been shown to exist, or the study of "beings, in so far as they are beings" (Cohen 2000).

Figure 2: An apple (the data) and its metadata attributes

Metadata can come in many forms, including library catalogues, nutrition labels, digital object descriptions, and (arguably) even memories[4]. For instance, the information displayed in a search result is a typical instance of metadata, detailing what a web document is (type) and its basic characteristics (title, link, time, and description) (Zeng and Qin [2000] 2016, 7). It follows that metadata can originate from a variety of sources. It can be generated automatically, manually entered, or a combination of both. It can be harvested through techniques such as web scraping, converted from other formats, or user-contributed (web publishing, file sharing, social media). The quantity and quality of what metadata describes and the standards to be withheld in its description is effectively limitless (Zeng and Qin [2000] 2016, 17). In the words of Admiral Grace Hopper, "The nice thing about standards is that there are so many of them

[4] *Object-Oriented Ontology* suggests that objects are represented by our preconceptions of their qualities, virtualities conceived by our consciousness (Bryant 2011)

to choose from" (Pomerantz 2015, 93); digital metadata exists in many formats and structures, with several organizations maintaining standards for their organization and usage.

Metadata is voluntarily provided as 'data exhaust',[5] a by-product of using digital technologies throughout our daily lives. The simple act of taking a photograph on a digital camera or smartphone automatically embeds information in Exif format to the image, detailing the parameters in which the photo was captured and often its GPS coordinates (if location tracking is enabled). A posted tweet on Twitter contains information cataloguing the time and location the tweet was published from and automatically monitors the number of interactions, likes, retweets and follows associated with it. Google Analytics (among others) makes a business of tracking data associated with a website and its key terms, including the traffic of users (geographically, numerically, chronologically, etc.). By simply typing in the URL of a website and opening it in a web browser, data has been collected and contributed towards the domain's data metrics. Methods of circumvention exist[6], but we have

---

[5] The term was purportedly coined by Josh Kopelman in a blog article (Kopelman 2008).

[6] Various web browser add-ons exist to monitor and inhibit trackers or offer an 'incognito mode' that reduces (but does not outright eliminate) the amount of data visible to tracking sites. You can even connect to the internet via a proxy to hide your location. Ultimately, the process of establishing a 'private' browsing experience is not rudimentary.

effectively "moved from a world that is “private-by-default, public-through-effort” to one that is “public-by-default, private-with-effort” (boyd, 2011).

2.1.1 The Defining Properties of Metadata

In his book *Metadata*, Jeffrey Pomerantz defines five different categories of metadata based on existing formats (Pomerantz 2015, 17–18):

1. *Descriptive Metadata*, which describes an object’s properties,
2. *Administrative Metadata*, which is concerned with how the object was produced, how the object can be maintained, and the legal usage restrictions in place (e.g. digitization, copyright),
3. *Structural Metadata*, which contains information on how an object is organized (e.g. chapters, layers),

Pomerantz also defines *Preservation Metadata*, which carries information on how to preserve an object, and *Use Metadata,* which describes how an object has been used (e.g. downloads and download details), although in my opinion these two categories could be considered subsets of administrative metadata[7].

[7] These definitions are not standardized; the IPTC has its own 3 main categories of data: administrative, descriptive, and rights (International Press Telecommunications Council n.d.). These definitions lack the preservation and use metadata descriptions defined by Pomerantz, and the administrative & rights definitions are separate in contrast to Pomerantz’s more conglomerated definition of administrative metadata. Despite the IPTC being a global standard, Pomerantz’s definitions appear more nuanced.

2.2 Data Visualization and Analytics

Data analytics is the process of determining useful judgements from information, and data visualization is the act of creating “visual representations of abstract data to amplify cognition”, often with an interactive computer aid (Meirelles 2013, 13; citing Card 1999). As this thesis is concerned with both the creation and *rendering* of contextual metadata, understanding the process behind data analytics and visualization is vital for any acceptable representation of the digital media I wish to investigate.

Joe Hammond and Paul White define three phases of data analysis:

- *Data acquisition*: the act of collecting data
- *Data processing*: the act of performing operations on data to convert it into other forms (generally for legibility purposes, ‘technical’ analysis)
- *Data interpretation*: the act of gaining insights from data (the analysis of processed data, ‘thoughtful’ analysis)

(Hammond and White 2015, 145)

As a subset of data analytics, data visualization incorporates all three elements in its process, as the act of creating visualizations of data is a manipulation of collected data into a simplified format for the visual communication of ideas defined by the visualization’s creator for the perceiver (Ware 2013, 4). For example, the process of creating an infographic on population density requires the acquisition of census data, which is then processed into a digital spreadsheet format with *summation* and *average* functions before finally being reviewed by a statistician who then produces a representative drawing showcasing key talking points for their institution. As Ware describes, “visualization applies vision research to practical problems of data analysis in much the same way as engineering applies physics to practical problems of building factories” (Ware 2013, 29). There is an implicit sensory bias in data visualization’s prevalence as a method of expressing data for interpretation, but the human visual system remains to be the highest bandwidth channel for information communication; “the 20 billion or so neurons of the brain devoted to analyzing visual information provide a pattern-finding mechanism that is a fundamental component in much of our cognitive activity. Improving cognitive systems often means optimizing the search for data and making it easier to see important patterns” (Ware 2013, 2).

Data visualizations have a number of objectives linked to cognition, many of which Meirelles summarizes as follows (Meirelles 2013, 13):

- to record information;
- to convey meaning;
- to increase working memory;
- to facilitate search;
- to facilitate discovery;
- to support perceptual inference;
- to enhance detection and recognition;
- to provide models of actual and theoretical worlds;

- to provide manipulation of data.

It is commonly understood that “working with the aid of thinking tools is much more cognitively powerful than that person alone with his or her thoughts” (Ware 2013, 2), and further “visual displays of information can be considered cognitive artifacts, in that they can complement and strengthen our mental abilities” (Meirelles 2013, 13). Visualizations act as one of such formats of cognitive aids, providing a visual, interactive layout for the “ communication, analysis, discovery, exploration, etc.” of information in a digital and/or material form (Cairo 2016, 28).

Visualizations and visible phenomena in general are composed of what Ware calls sensory and arbitrary symbols: “sensory is used to refer to symbols and aspects of visualizations that derive their expressive power from their ability to use the perceptual processing power of the brain without learning. The word arbitrary is used to define aspects of representation that must be learned, because the representations have no perceptual basis” (Ware 2013, 9), and elaborated, “sensory representations are effective because they are well matched to the first stages of neural processing. They tend to be stable across individuals, cultures, and time. […] Conversely, arbitrary conventions derive their power from culture and are therefore dependent on the particular cultural milieu of an individual.” (Ware 2013, 9). For example, the scientific sign for electrical resistance, Ω (uppercase omega) would be utterly nonsensical to anyone unfamiliar with the nomenclature, but an ellipsis drawn around a series of objects intuitively defines a border between a space within its boundaries (the space holding the symbols) and the space external to its shape (Ware 2013, 7).

As a less abstract example, the act of observing a still art scene benefits from cognitively recognizing the forms, colours and existence of fruits. The viewer's understanding of the scene is built upon both the foreknowledge of the existence of this food-matter (arbitrary) as well as the capacity to convert visual sensations into an understanding of the image's context (sensory). Imagine the fascination one might have simply seeing a rendered image (photo, painting, or otherwise 'realistic') for the first time in one's life. Without some degree of cognitive processing our visual field may as well be a disorganized mass of colour splotches (assuming we can even recognize 'colour' or 'splotches' in such a scenario). Furthermore, our apparent conscious experience is unique; what I observe at any given moment will never be more than an imitation of another. We can experience empathy and take measures to emulate an experience, but the complexities of our universe make duplication improbable, if not impossible. Although I may simulate the experience of being colourblind with technical aids, I will never have the lived experience of a colourblind person, never mind the conscious experience of another human being. The point of this digression is to emphasize the importance and challenge faced in analytics and visualization of conveying meaning from datasets and analysis into cognitively digestible information or cognitive aids for thought and exploration.

2.2.1 Geospatial Representations

Maps are effectively visualizations of the spatial distribution of data, generally concerning defined areas or geographical regions. According to Isabel Meirelles, spatial-temporal phenomena can be organized into three main types (2013, 161):

1. Existential changes, referring to instantaneous events and the appearance and disappearance of objects and/or relationships.
2. Spatial changes, referring to transformations of objects such as location, size or shape.
3. Thematic changes, referring to changes in values or attributes of space (e.g. demographic maps).

These 'events' summarize the content that can be projected on a map for the purpose of data visualization. While imperfect, mappings provide guidance in describing how a network or area can be navigated; in spite of its high resolution graphics and enormity, *Google Earth* is but a representation comprised of patchwork snapshots of the planet, each taken at non-simultaneous periods of time (Farman 2010, 874)[8]. Despite this, most of our representations of the planet strive for a visual consistency in lighting, spatial distribution, political borders and other subjective delineations (an oxymoron). Farman suggests that "the view that maps are objective visualization tools is conceived out of a larger misunderstanding of space and territory," referring to H. Lefebvre's statement that "'we know that space is not a pre-existing void, endowed with formal properties alone. To criticize and reject absolute space is simply to refuse a particular representation, that of a container waiting to be filled by content'"[9] (Farman 2010, 874).

On May 1st, 2000, the United States granted civilian access to the signals transmitted from its Global Positioning System (GPS), a group of 24-29 satellites that beam radio signals to the earth's surface which can be utilized by a receiver to determine its position on the planet (Letham 2008, 5). As Farman notes, this brought to fruition "the much-speculated age of ubiquitous computing by

---

[8] "One reason that mapping technologies such as *Google Earth* often avoid critique is their use of satellite and aerial photography. Though the photograph has undergone intense scrutiny in the digital age in regards to its status as an index of reality, the photograph still holds a connection to material space that is unmatched by hand-crafted maps. Peirce (1998a: 322), who famously wrote that 'representations have power to cause real facts', brought notions of indexicality in visual representations to the forefront of semiotics." (Farman 2010, 874) referencing Peirce, C.S. (1998b) 'Of Reasoning in General', in Peirce Edition Project (ed.) The Essential Peirce: Selected Philosophical Writings Vol.2, pp. 11–26. Bloomington: Indiana University Press.

[9] from Lefebvre, H. *The Production of Space*. Translated by D. Nicholson-Smith. Oxford: Blackwell Publishing, 1991.

moving the user interface away from the personal computer to the space of pervasive computing" (2009, 1). This enabled the device users the capacity to know almost precisely where they are on the planet at any given time when connected to the network as well as an improved ability to collect locative information about a space through networked technologies. Further, this locative information could be aligned with existing mappings and representations of spatial data to provide more informative data visualizations related to place, the critical element that drives the inquiry of this thesis.

Discussing the cognitive benefits of spatial visualizations, Meirelles writes "we process spatial properties (position and size) separately from object properties (such as shape, color, texture, etc.). Furthermore, position in space and time has a dominant role in perceptual organization, as well as in memory" (Meirelles 2013, 19) and "spatial encoding is central to how we construct visualizations, in that the geometric properties and spatial relations in the representation—the topology—will stand for properties and relationships in the source domain." (Meirelles 2013, 20). This establishment of locative parallels between visualizations and context allows us to communicate the spatial qualities of these visualizations more effectively to users.

2.3 Digital Photography and the Digital Image

Digital photographs are binary representations of images, data to be translated into visual projections on our screens or print media. But digital photographs as a *thing* are effectively machine instructions for graphical representation, as opposed to the visual stimulus that we often associate with them. Citing George Legrady, Rosen indicates that the digital photograph is a mere 'simulated photographic representation' created "according to a certain range of prior pictorial norms … identified with photography"(Rosen 2001, 308), referring to the 'machine instruction' nature of code that transforms data-representations of images into visual stimuli. This nature of the digital image allows for it to carry along embedded information, often in the format of metadata. Still, the digital photograph's roots in analog photography are unmistakable: from its glass lenses to the representations grafted into silicon memory, digital photography hardware and terminology are largely used in a traditional manner. Whereas the earliest photographs were emblazoned into metal plates, modern photographs are easy to capture, store, accumulate, exchange, and carry, as digital media is as lightweight as the physical device storing the information[10].

[10] with cloud computing, such a file may not even be physically present with the user

2.3.1 Dissecting Digital Graphics

With computer technology, images have become “a pliable sequence of digital data and electronic impulses” (Batchen 2001, 179), effectively raw data organized by software algorithms to be perceivable by a rendering application[11]. Although they are distinct in their makeup, traditional photography and digital photography practices have a common need of subject matter, processing, and visual output:

> “The photograph is now a type of ‘algorithmic image’; a term we use in order to indicate that the image has to be considered as a kind of program, a process expressed as, with, in or through software. When the photograph became digital information, it not only became malleable and non-indexical, it became computational and programmable.” (Rubinstein and Sluis 2013, 29)

When created, digital photographs usually contain metadata about where and when they were captured, the option settings of the camera, as well as the photographic image itself, all generated at the moment of capture (Pomerantz 2015, 96). Most modern cameras use Exchangeable image file format (Exif) metadata schema, which describes the camera and the settings and conditions in which it was shot. The website *I Know Where Your Cat Lives*[12] utilizes the Exif

---

[11] Daniel Rubinstein and Katrina Sluis note that “The same data could be just as easily output as a text file, a sound, a string of numbers or remain unprocessed.” (2013, 27)
[12] https://iknowwhereyourcatlives.com/

metadata of one million public images of cats and positions them in on a world map (see *figure 3*) (Mundy et al. n.d.).

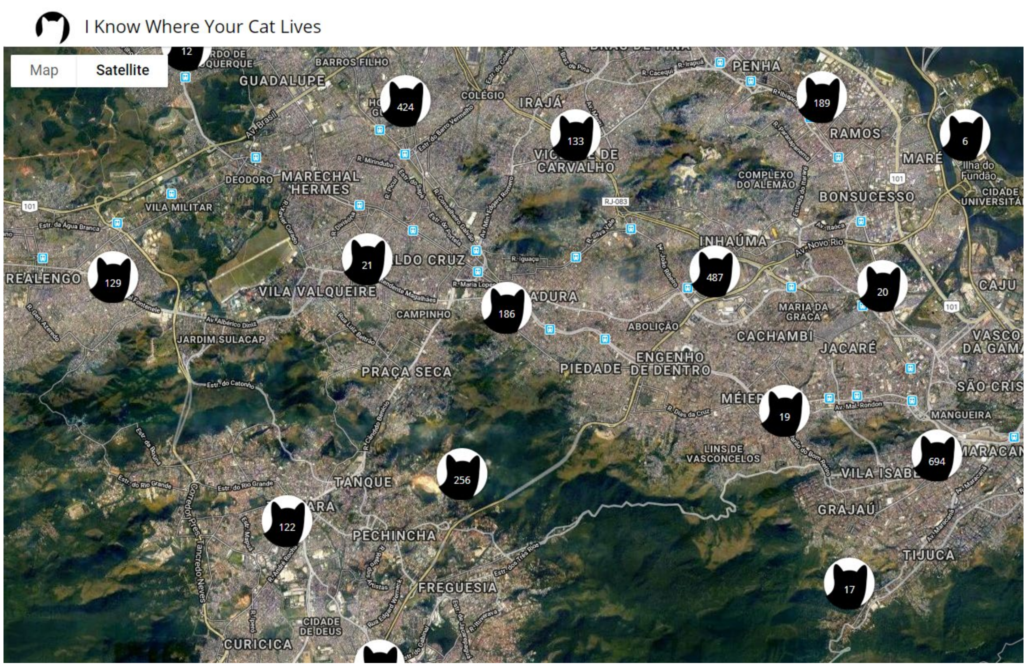


Figure 3: Screenshot of *I Know Where Your Cat Lives* (Rio de Janeiro, Brazil)[13]

2.3.2 Metadata Retention on Social Networks

In 2019, it was reported that Instagram and Facebook held over 1 billion monthly active users respectively (Noyes 2019; Statista 2019), with tens to hundreds of millions of images uploaded each day (Desjardins 2017; Noyes 2019). Another report indicated half a billion tweets were produced each day by Twitter's 336 million active monthly users (Lin 2019; Iqbal 2018). With the sheer volume of content uploaded by the minute, it begs the question how that media is being

[13] *I Know Where Your Cat Lives* by Mundy, Owen et al., Accessed November 24th, 2019: https://iknowwhereyourcatlives.com/

handled and stored within these social networks. In 2017, Flickr backend engineer Archie Russell blogged that on a high traffic day the image hosting site receives as many as 25 million photo uploads. In order to help manage this, Flickr uses compression algorithms to reduce the size of the uploaded images with minimal loss in quality: "these tools work by decoding the JPG, then very carefully compressing it using a more efficient approach. This typically shrinks a JPG by about 22%. At Flickr's scale, this is significant" (Russell 2017).

The website *Embedded Metadata Manifesto* lists many social media sites and systems and indicates whether the metadata of the original images uploaded are preserved correctly through the process of uploading to downloading. Facebook, Instagram, Twitter, LinkedIn, and Tumblr number among these platforms; Tumblr does maintain Exif [14]metadata format in specific circumstances but otherwise metadata embedded into images, upon upload to these social platforms, is removed and ultimately irretrievable (International Press Telecommunications Council 2015). Despite claims to retain its context "when travelling across software, devices and databases" (Adobe Systems, Inc.

[14] Exif is a standard metadata format for photographic images (including digital rights management). See

n.d.), extensible metadata formats such as Adobe's XMP are still vulnerable to removal from their media[15].

It is technically possible to add any sort of metadata tag to common exchangeable media formats, such as picture files, videos, sounds, etc. but most prominent social media networks scrape the bulk of this metadata away before it can be shared on their platform. The resulting uploaded file no longer contains details about its origin and may even receive injected metadata content (as is the case with Facebook). This is of course problematic to the interests of injecting additional contextual information to images as it limits the ability for metadata information to be distributed on social networks[16].

2.4 The Mobile Interface and Hybrid Spaces

Mobile Interface Theory is a term used by Jason Farman to describe the relationship between people and the digital devices they carry around in their daily lives. The mobile interface essentially acts as an extension of the human

[15] XMP is an open source extensible metadata format that gives users the capacity to add arbitrary metadata elements to their media. XMP stands for eXtensible Metadata Platform and is an ISO standard for metadata. If the media does not support XMP, the XMP information is stored in a separate 'sidecar' file.

[16] My 'guerilla' solution to this, steganography, is discussed in Section 4.2.1

experience, the interface between an observer and their world; we are effectively cybernetically-enhanced by the devices we carry, providing us windows to virtual planes of existence (Weibel 2009, 226). Although consumer-targeted mobile devices provide a 'lower-fidelity' connection with virtual spaces than 'immersive' non-portable headsets and studio gadgets, they are ubiquitous and accessible technologies. Created in 2000, the activity of *Geocaching* utilizes the mobile device as a means of navigating the physical 'play space' of reality. Players can search for one of hidden 'caches' using a series of clues provided by other participants through acts of "asynchronous documentation of presence" (Farman 2009, 4); unlike many geospatial games that feature elements such as shared presence, players participate at different moments in time. Elizabeth Grosz believes that the phenomena that we perceive as reality is defined by the images created by our sensory processes. As Grosz describes, "the body phantom is the condition of the subject's capacity not only to adapt to but also to become integrated with various objects, instruments, tools, and machines. It is the condition of the body's inherent openness and pliability to and in its social context" (Grosz 2001, 35). This parallels the Deleuzian assertion that "every actual surrounds itself with a cloud of virtual images''[17] (Deleuze and Parnet [1977] 2007, 148) which include the images we perceive in our reality as well as

[17] Levi R. Bryant argues that these virtual representations are "always the virtuality of the substance", observing Deleuze's tendency to treat the virtual as separate from the substance (Bryant 2011, 30).

representations displayed by our devices of virtual realities[18]. The virtual space becomes another dimension to our existence, and our relationships with mobile interfaces allows for virtual spaces to utilize reality like a cartesian plane marked by the vectors of our mobility and interactions.

A 2013 report from the Pew Research Center states that 45.9% of the American population[19] owns a smartphone and 75% own a web-enabled device that doubles as a mobile phone[20] (Thompson, Purcell, and Raine 2013). Farman's investigations[21] into the prevalent phenomena emergent in our portable connections to virtual environments[22] and its hybridization with reality are of critical importance in this thesis investigation. Core to Farman's notion of Mobile Interface Theory are *hybrid spaces*, which

> "are formed through a combination of three elements: social interaction, digital information, and physical space. The digital information people access in hybrid spaces is not exterior to the

---

[18] Also: "Reality has always been interpreted through the reports given by images; and philosophers since Plato have tried to loosen our dependence on images by evoking the standard of an image-free way of apprehending the real. But when, in the mid-nineteenth century, the standard finally seemed attainable, the retreat of old religious and scientific thinking did the contrary, the new age of unbelief strengthened the allegiance to images. The credence that could no longer be given to realities understood in the form of images was now being given to realities understood to be images, illusions."(Sontag [1977] 1990, 153)

[19] Also, "by 2021, 40 percent of the world's population is projected to own a smartphone." (Holst 2018)

[20] Worth noting: approximately 60% of Americans own a Laptop, and 51% access the internet using a mobile phone (Thompson, Purcell, and Raine 2013)

[21] in spite of the important investigations covered in his material, it should be noted that Farman arguably does not establish a theory with the rigour necessary to be considered an academically justified theory (based on his writings in *Mobile Interface Theory*) *(Crooks 2013, 3)*

[22] By *virtual environments* I am referring to non-real spaces that we can engage with

> place; it becomes a part of that place for the user, just as a street sign or other physical informational becomes a part of a place. Hybrid space is a valuable conceptual tool because it refuses the urge to separate location-based digital information from the physical place it describes. Instead, the digital plays a role in shaping how people 'read' physical places." (Frith 2015, 8)

Mobile interfaces reduce the distance of communication, establish their own rules of engagement and function, and alter how we interact with our own reality. The virtual interface could be as simple as a text messaging service or as complex as Niantic's *Pokémon GO* (2016), but both effectively hybridize our physical lives with a virtual existence and communication medium[23]. As new media becomes part of the everyday environment and our social media profiles become part of our identities (Arvidsson and Delfanti 2019, 48–49); we are becoming so immersed in new media that we have difficulty being perceptive of the encompassing effect it has on our lives (Farman 2014, 49; Rohlinger 2019, 1). By simply interacting with and carrying these devices with us on a regular basis, we share contextual data exhaust as dictated by our behaviours and the functionality of our device's features. It is these intersections of our daily physical lives and the virtual amazon that I hope to investigate in mobile device usage. The development of a mobile application for the capture, manipulation,

[23] These communications are colloquially called *new media*, "the mass communications that rely on digital technologies such as social media, online games and applications, multimedia, productivity applications, cloud computing, interoperable systems and mobile devices" (Rohlinger 2019, 1).

and rendering of contextual data will assist in my investigation of how the external context associated with photographs can augment the hybridized, digital experience.

3.0 Methodology / Research Process

This chapter discusses how the three methodologies that I have selected apply to my research and how I intend to incorporate them into the research process. I will also outline the method that I will use to assess my Discursive Design research: use cases and user groups. The three methodologies are as follows:

1. *Research-Creation*, where creation and research in the design process complement and support one another through the process of developing an artefact
2. *Discursive Design*, a design approach involving the creation of an artifact that serves as a vehicle for discussion concerning the artifacts existence in relation to existing socio-cultural values
3. and *Critical Design,* a methodology concerned with the critical analysis of the ethical and ideological context of a design.

3.1 Research-Creation

In "Research-Creation: Intervention, Analysis and Family Resemblances" (published in 2012), Owen Chapman and Kim Sawchuk briefly outline "four different modalities in which research and creation are linked within current academic practices …":

1. “Research-for-creation,” the gathering of materials, practices, technologies, collaborators, narratives, and theoretical frames that characterizes initial stages of creative work and occurs iteratively throughout a project.
2. “Research-from-creation,” the extrapolation of theoretical, methodological, ethnographic, or other insights from creative processes, which are then looped back into the project that generated them.
3. “Creative presentations of research,” a reference to alternative forms of research dissemination and knowledge mobilization linked to such projects.
4. “Creation-as-research,” which draws from all categories, an engagement with the ontological question of what constitutes research in order to make space for creative material and process-focused research-outcomes.
(Chapman and Sawchuk 2015, 49)

The research conducted in this thesis can be considered a conversation between the first three components of Chapman and Sawchuk’s modalities[24]. To thoroughly understand the existing research and context surrounding the topics discussed in my thesis, it was naturally important to conduct research on the subject matter. The research provided supporting justification for my arguments and assisted in my understanding of how a prototype might be designed and assembled in a manner that addresses the research questions at hand. This research led to the creation of a proof-of-concept prototype assuring that the topic was well within my capacity to investigate but accomplishing the double-duty of providing insights that looped back into the research process. The final

[24] The fourth, *creation-as-research*, can be presupposed to be an element by nature of this thesis utilizing the other three modalities; I will leave justifying the methodology to scholars concerned with the subject.

prototype serves as a creative output for the thesis investigation, responding to my expectations of the future of networked technologies, the process contributing to the Research-Creation feedback loop and as a discussion piece for developers concerning how future innovative prototypes employing contextual metadata can be better developed.

Research-for-Creation and *Research-from-Creation* have been used rather interchangeably with *Research-for-Design* (RfD) and *Research-through-Design* (RtD) terminology, depending on the discipline and the scholars discussing the matter. Case in point, the chapter "Research through Design" from *The Encyclopedia of Human-Computer Interaction* outlines both RtD, RfD, and related processes as they are investigated by various external authors (Stappers and Giaccardi 2014). Thorough examples outlining the process of carrying out these research methodologies are used to reinforce the Stappers and Giaccardi arguments on the subject. Stappers and Giaccardi establish that *research* aims for general knowledge in an abstracted form with a long-term focus with the establishment of a theory as its outcome, whereas *design* is often focused on a specific solution situated within a current project, the outcome being the realizations made through the process of design (2014).

3.2 Discursive Design (and Speculative Design)

While Research-Creation outlines the process in which the research will be conducted, it provides little insight into how I will come to conclusions about the topics discussed and ultimately provide resolution and a benchmark to evaluate my research questions. Because the domain in which my investigation is framed is inherently concerning a future that is yet-to-be and defined upon subjective notions of contextuality, I have conceded that it is beyond the scope of this investigation to determine any sort of optimal strategies or conclusions. Instead, I seek to propose and present a functional, contextual metadata framework to facilitate discussion on how such a format might become integrated with existing technological infrastructures and the potential target markets for its associated products. As a means of investigating these questions and possibilities, I seek to employ Bruce and Stephanie Tharp's Discursive Design methodology as well as ideas from Auger's Speculative Design.

Discursive Design is a design approach coined by Bruce and Stephanie Tharp that broadly encompasses a family of design methodologies such as Speculative Design and Critical Design. The primary design intention of Discursive Design "is not utilitarian in the typical sense but rather to communicate particular ideas and rouse reflection," where "the material language, traditions, and characteristics of design are employed for immaterial aims" (2019, 7). In other words, 'discursive objects' are created as works facilitating discussion on particular topics of interest[25]. The discursive object is designed with the intention to be appraised "as the deliberate consequence of some set of sociocultural values and ask what does it mean or say about culture that this product exists or could exist [...] better magnifying, reflecting, and revealing aspects of culture for its audience [or] intentionally distorting in order to emphasize, propose, speculate, instigate, or criticize" (Tharp and Tharp 2019, 13).

It is worth noting that Auger intended for Speculative Design (rather than Discursive Design) to be the encapsulating methodology. Referencing other members of this family of speculative/discursive methodologies: "as one of the core motivations of this practice is to shift the discussion on technology beyond the fields of experts to a broad popular audience, the choice of 'speculative' is preferable as it suggests a direct correlation between 'here and now' and existence of the design concept"[26] (Auger 2013, 2). In an interview on Croatian

---

[25] Tharp and Tharp refer to two basic conceptions of discourse: the commonly understood adjective for "communicative acts and exchanges" (i.e. discussions), and Michel Foucault's notion of discourse as a system of thought and knowledge. (Tharp and Tharp 2019, 74–76)

[26] "the word 'fiction' before design immediately informs the viewer that the object is not real; 'probes' infer that the object is part of an investigation; and both 'discursive' and 'critical' reveal the intentions of the object as an instigator of debate or philosophical analysis. These terms act to dislocate the object from everyday life, exposing their fictional or academic status. For those within the design or research community these semantic details are less problematic as

website *Speculative*, James Auger outlines the key themes of Speculative Design in place of a formal definition[27]:

> At this time there are three basic themes: (i) Arrange emerging (not yet available) technological "elements" to hypothesise future products and artefacts, or (ii) Apply alternative plans, motivations, or ideologies to those currently driving technological development in order to facilitate new arrangements of existing elements, and (iii) Develop new perspectives on big systems. With the purpose of: (i) Asking "what is a better future (or present)?" (ii) Generating a better understanding of the potential implications of a specific (disruptive) technology in various contexts and on multiple scales – with a particular focus on everyday life. (iii) Moving design "upstream" – to not simply package technology at the end of the technological journey but to impact and influence that journey from its genesis (Auger 2016).

When I initially began my research my intention was to cover all of the above themes and criteria, but I realized through the development of my 'proof-of-concept' prototype that the scenario I expected to take place in an alternate present or future could just as well be society as it exists in the present. However, the definition efficiently encapsulates the core objectives of my research investigations and Auger's term 'alternate present' is efficient for comparing our present experience to another impacted by a ubiquitous external

---

familiarity with the discourse makes the terminology less important, but for those unfamiliar with these practices, semantics fundamentally affect how the work is experienced and assessed." (Auger 2013, 2)

[27] Auger states that his opinion of the Speculative Design definition "changes daily" (Auger 2016), noting the temporal, methodological, and diverse qualities that make each Speculative Design project unique challenge rigorous definitions (Auger 2013, 21). Likewise, Tharp and Tharp indicate that "we expect that the structure will be adorned, adjusted, and added onto as the field matures." (2019, 74)

contextual metadata framework (which reflects Auger's intentions with the term); I will continue to use this term for the remainder of the thesis to illustrate this.

While I appreciate the thoroughness of Auger's Speculative Design and related writings, I find myself obliged to use the term Discursive Design to describe my methodology as it comes across to me as a more succinct means of describing the objectives of my prototyping process and the succeeding critical discussion regarding my theoretical multimodal framework and the prototyping process, without the distractions of the terms 'speculative', 'speculative futures', and the bias invoked by emphasizing alternative timelines and scenarios.

3.2.1 Critical Design

Considered a subset of both Discursive and Speculative Design methodologies by Auger and the Tharps, Critical Design was coined by Anthony Dunne and Fiona Raby in the mid-nineties (2013, 34) and the definition has since faced criticism for its broad title[28] and unclear definition[29]. To address this, *Speculative*

[28] Dunne and Raby actually admit in *Speculative Everything* that the name they chose ('Critical Design') was not a clear representation of their methodology, but rather one chosen to create controversy and discussion: "naming it critical design was simply a useful way of making this activity more visible and subject to discussion and debate" (2013, 34)

[29] e.g., "Designers start by identifying shortcomings in the thing they are redesigning and offer a better version. Critical design applies this to larger more complex issues. Critical design is critical

Everything provides several clarifications and an updated definition for Critical Design: "By acting on peoples' imaginations rather than the material world, Critical Design aims to challenge how people think about everyday life. In doing this, it strives to keep alive other possibilities by providing a counterpoint to the world around us and encouraging us to see that everyday life could be different" (Dunne and Raby 2013, 45).

The same year, Shadowen and Jeffrey Bardzell spoke on a panel titled "What is Critical About Critical Design?", analyzing the Dunne and Raby definition and addressing what they considered to be glaring issues with the methodology. The Bardzells proposed an alternative 'reconstituted' definition, with Human-Computer Interaction (HCI) research in mind: "critical design is a design research practice that foregrounds the ethical positioning of designers; this practice is suspicious of the potential for hidden ideologies that can harm the public; it optimistically seeks out, tries out, and disseminates new design values; it seeks to cultivate critical awareness in designers and consumers alike in, by means of, and through designs; it views this activity as democratically participatory" (Bardzell and Bardzell 2013). The Bardzells succeed in using precise language to frame Critical Design within more clearly delineated goalposts, referring over the

thought translated into materiality" (Dunne and Raby 2013, 35); their passages tend to read as more of a manifesto.

past 150 years of critical theory and thought to support their position. In contrast, Dunne and Raby reject the influence of critical theory, the Frankfurt School, or 'plain criticism' but rather emphasize the importance of not making assumptions (2013, 35); stating "A critical design should be demanding, challenging, and if it is going to raise awareness, do so for issues that are not already well known. Safe ideas will not linger in people's minds or challenge prevailing views but if it is too weird, it will be dismissed as art, and if too normal, it will be effortlessly assimilated" (2013, 43).

Critical Design is not the primary methodology I intend to incorporate into my thesis work, but it provokes an investigation into the thesis' ramifications on society and culture. My thesis prompts a multiplitude of questions concerning ethics, ideologies and values in its intersection with existing infrastructures of social communications, data transmission and recreations of archived experiences. Any metadata framework encapsulating contextual data may impede upon privacy and existing social frameworks of communication, impact how we capture our experiences, our expectations in our immersion in media, and it may encourage dialogues discussing the differences between the past and present through archived media. These are just a few examples of how Critical Design can play into my research concerns, and the understanding of Critical

Design from both groups of researchers described above will fundamentally support my approach to this investigation.

3.2.2 Use Cases and User Groups (a Method for Discourse)

According to Elizabeth and Richard Larson, a use case describes the manners in which an end-user (or user group) would want to use a system and consist of three core elements:

1. The System (as framed by the boundaries of the application),
2. Actors (or User Groups, external entities that interact with a system),
3. and Use Cases (processes occurring within the system)
   (Larson and Larson 2004) .

For the purposes of conducting my Discursive Design analysis, I will be discussing the various roles involved with multimodal frames based on such use cases and their user groups. Each user group will represent users with shared objectives and values in relation to contextual metadata as embedded through multimodal frames (see Chapter 4) to better frame critical discussion on how these multimodal frames may be used in practice.

The ubiquitous incorporation of a contextual metadata framework (Chapter 4) in modern technology extends to an innumerable quantity of possible projects that could utilize the concept and the multiple market groups involved with these projects from life to death; the use case system is broad in scale. Accurately establishing representative characters or groups of characters (i.e. the UI/UX Persona method[30]) would be difficult, but groups of occupations still emerge: photographers, consumers, archivists, and entities of surveillance. I am primarily concerned with how the intentions and interactions of these various entity groups involved with contextual metadata technologies will manifest from an abstract standpoint rather than concerning ourselves with how individuals, small groups, or automated software instances might react to the experience and effects of contextual metadata's emergence.

This method will be discussed in the Use Case Discussion in Chapter 5, where I introduce several User Groups involved with the concept of multimodal capture and ECM (to be explained in the forthcoming chapter).

[30] Worth noting, I initially began working with a modified version of the Group Personas method before coming to realize that Use Cases more accurately described the type of analysis I was engaging with in Section 5.0, hence their 'persona'-like appearance.

4.0 Theoretical Digital Framework

During my research, I have reached the conclusion that there is a niche in metadata standards that remains largely unexplored, particularly for users on the consumer side of information technologies. As Dey and Abowd state in their journal on Context and Context-Awareness, “By improving the computer’s access to context, we increase the richness of communication in human-computer interaction and make it possible to produce more useful computational services” (Dey and Abowd 2000, 1). As stated in the Introduction, I seek to propose a new framework for digital media metadata that integrates a description of a digital media object’s context that I dub *external contextual metadata* (ECM): metadata that describes information about the context of an object that are external to the object’s inherent properties.

This chapter is dedicated to outlining my thoughts concerning how ECM could be integrated into networks and ubiquitous technologies as well as how I intend to integrate it within my research. My objective as the author of this thesis is to demonstrate the potential usefulness of ECM. As evidenced by the scope of my context / literature review, my research is implicitly biased towards embedding such information into digital photography related technologies. The contextual metadata framework incorporating ECM that I will introduce in this chapter is designed to facilitate research and provoke discussion on the matter rather than to establish a new standard, which is well beyond the scope of a Master's thesis.

I will be frequently using the term *renderer* or *renderer application* to describe any application designed to convert data and metadata into a sensory representation for the remainder of this thesis. This is similar to Lev Manovich's idea of computer metamedia, which is used by software to simulate "prior physical media" or produce new computer objects without physical precedent (Manovich 2013, 110); renderers could be considered a subset that specifically appeals to software that can produce perceivable information. Given the emphasis on digital photography and data visualization thus far, the use of the term will largely reference renderers that produce visual representations.

4.1 External Contextual Metadata (ECM)

Discussing photography, Barret describes three categories of pictorial context:

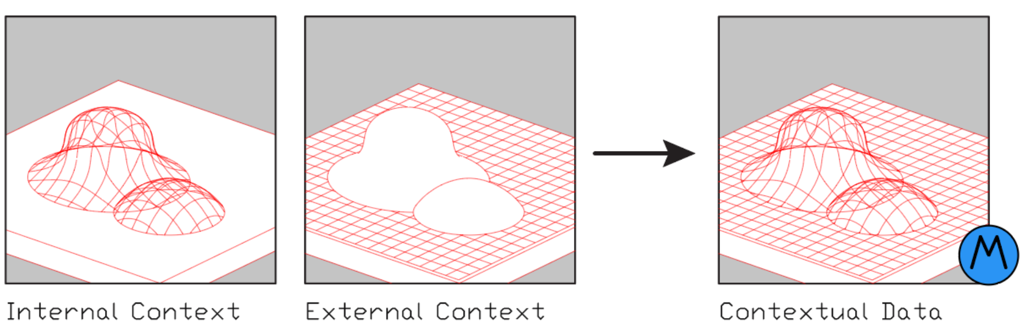
1. <u>Internal:</u> what is descriptively evident; subject matter, medium, form, relations among these three. Content that can be understood about a photograph without prior knowledge of its contents
2. <u>Original:</u> Content in an image that requires knowledgeable understanding of the reasoning behind the capture and general background concerning

deliberate meaning behind the internal elements (e.g. a wartime photo has more meaning with knowledge of the war). Includes psychological, social, historical (personal and otherwise) context

3. External: the situation in which a photograph is presented or found (e.g. a gallery, billboard, game) (Barrett 1999)

A photograph can be considered a container of (visual) information. Looking back to Ware's descriptions of *sensory* and *arbitrary* visualization attributes (Data Visualization, Section 2.2), we might relate *Internal* context to sensory attributes and *Original* context to arbitrary attributes. What remains is the notion of External context, or the situation in which the container is positioned. As an element of this external container, photography cannot be considered autonomous of its rules and relationships with its two-dimensional plane that bears the processed image of binary data (a parallel to the analog 'silver halide'). Effectively, all images exist in relation with an object and the processes of rendering the image visible (Batchen 2001, 109).

Referring back to Pomerantz's five definitions of metadata (see Defining Properties of Metadata, Section 2.1.1), if we consider the properties of the context external to an object, we could propose a sixth category of metadata which we can call *External*[31] *Contextual Metadata*, metadata that describes the contextual background of an object. This could include information such as the population density of the area in which the object is located, or the average ambient temperature during the period of its creation, provided that such details are not inherently qualities that describe the object itself or otherwise fall into Pomerantz's five other categories of metadata. Since these properties concern how an object's properties position itself within the context in which it was created or presently exists, they would describe a layer of information that could provide interesting information about its relationships with external, otherwise unrelated elements. Through the automatized attachment of this information to an object's metadata in conjunction with a description of their relationship (e.g. the weather *at this object's position)*, an indexical connection is formed between the ECM and the digital object; otherwise, there may be no meaningful association between the object and the data (Braun [2001] 2015).



[31] 'External' is necessary as an adjective to distinguish from other forms of context, since internal context is effectively already described by metadata standards per Pomerantz's definition.

Figure 4: The relationship between Internal and External Context

This information is currently available and capable of being injected into several existing metadata standards. Although ECM has been collected for marketing applications for years (see Section 5.6: Surveillance Networks), up until this point there has not been a publicly available standardized metadata format that acknowledges ECM in this manner. The closest comparison to be drawn would be to the 'arbitrary data' embedded in XMP (Adobe Systems, Inc. n.d.) and other extensible metadata formats (which are deliberately undefined). There are also numerous software applications that provide contextual information (possibly in some structured format) but they are not publicly standardized metadata resources for multipurpose usage[32], nor explicitly understood to be 'external contextual metadata'.

4.1.1 Rendering Applications

For a digital image to be viewed, the data describing it needs to be translated, *rendered*, to a visual medium perceivable by an image-viewer's user. It follows that for the data embedded in a multimodal frame to be examined, it needs to

[32] It is not within the scope of this thesis to analyze the methods used by comparable programs to store contextual data, but data analytics programs certainly exist that have access to contextual datasets. These formats, however, are not exchanged by consumers.

be described in a perceptible context, whether it be through data visualization methods, abstract visualizations, or further obscurity[33]. How this data can be represented and what data is selected to be represented is left to the developers of rendering applications.

Representations of data can have varying levels of faithfulness to the source material as well as varying levels of success at accomplishing the intended task. According to Peircean semiotics, these representational objects can have any combination of three characteristic semiotic elements: icons (semblance of objects), indexicals (relationships corresponding to objects) and/or symbols (conventional stand-in for an object) (Atkin 2013). For simplicity's sake, we can visualize the varying levels of faithfulness of a representation as a spectrum of [literal] - [figurative] - [non-representation]. Literal representations would be those that strive for the most accurate representation of their data (e.g. data visualizations for business analytics). Figurative representations remain representations but are less concerned with accurately portraying data in favour of other design intentions (e.g. theatrical symbolism, metaphors). Non-representation is, quite literally, where a representation *almost* ceases to represent its data with any accuracy because the semiotic elements of the

[33] Although I use predominantly visual language here, I would like to point out that these representations could just as well be multi-sensory

representation have lost their understood connection to the original object (e.g. Glitch Studies, Section 6.1.3). However, it should be noted that icons, indexicals, and symbols rely on an understanding of what their conventions represent; a person uninformed of the communicative intention of the symbol or unaware of the indexical's relationship to the target object will observe it to be meaningless as a representation. Should the semiotic elements fail to demonstrate their intended meaning, a representation will inevitably tend towards the right side of the spectrum presented.

4.1.2 Mobile Metric Data

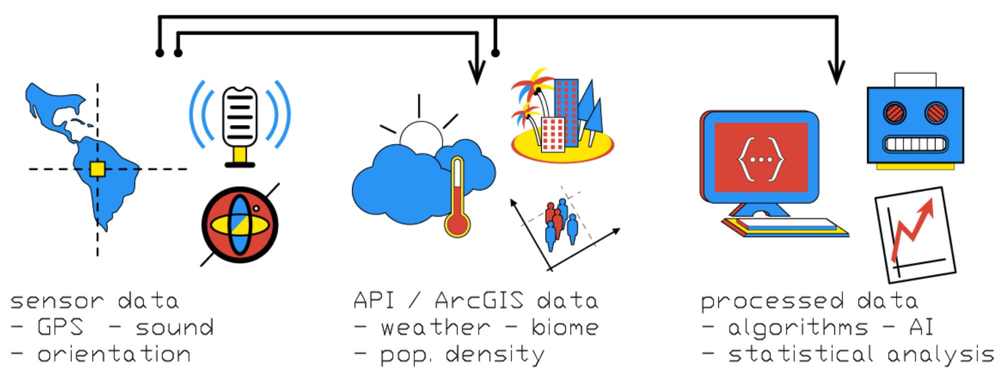


Figure 5: Accessible Data for Smart Devices

Metric data are data that describe some sort of measurement. There are several potential sources for metric data for smartphone devices. Smartphones have access to a variety of services and features that can be tapped as a source of metric data. These metric data sources can be roughly broken down into three categories (see *figure 5*).

1. *Sensor Data*: any metric that can be collected by *dedicated* sensors built-in to a device, such as sound, orientation, and light intensity.
2. *Reference Services*: any metric provided by external services or mappings, such as weather APIs or ArcGIS mappings.
3. *Processed Data*: any metric acquired from the analysis of existing data, such as machine vision algorithms and statistical analysis.

The word *dedicated* is emphasized as otherwise, any Sensor Data could be misunderstood to be Processed Data since captured readings must be signal processed in order to be digitally legible. One notable outlier, GPS, usually requires specialized sensors to capture as well as the satellite-based service signals that broadcast it; I would consider this Sensor Data as a device's GPS sensors are typically dedicated to this function. Some API's also provide Processed Data as a service in exchange for input data, such as IBM's Watson. Regardless, the intention here is to provide a rough outline of three key sources of data that a mobile phone has access to which are relevant for the forthcoming discussion.

If we consider data collected from sensory sources as 'captured' or 'primary-source' metadata and data processed or looked up from a reference service as 'annotative' or 'secondary-source' metadata, we can distinctly identify that this secondary-source metadata relies on primary-source metadata because it requires some sort of indexical relationship between itself and an anchoring contextual element (e.g. a location, a sound, or a picture compared to their secondary descriptions). Secondary-source metadata can also be collected and annotated to primary-source metadata after a moment of capture, since it consists almost entirely of database references and analysis.

4.1.3 Flaws and Proposals for ECM

Given the incredible abundance of circumstances that could surround an object it is easy to dismiss ECM on the basis that the data that defines these contextual parameters are difficult to define, standardize, or in many circumstances, collect. There is also the matter of the origin of the contextual data: how was it measured, how was it collected, what parameters are in place and what biases may affect its definitions? In the previous section, we also established a key challenge concerning the *need* for appending ECM to digital objects, as it may largely consist of this annotative variety of metadata; why append ECM to an object when it can be selectively looked up?

In response, I propose the creation of a *contextual* metadata framework unifying 'internal' contextual metadata, metadata as it exists presently, with 'external' contextual metadata. The following considerations were pulled together after researching various methods of storing and organizing data in digital objects and subsequently speculating how ECM might be integrated in such a format (with particular reference to Pomerantz 2015; as well as Zeng and Qin [2000] 2016). The list is by no means exhaustive, but rather acts as a series of proposals towards the development of a standard that could exist in the future.

1. *Contextual metadata can be stored in a JSON-style or similar format*

   Already successful in the storage and transmission of data for APIs, the JSON format can provide a hierarchical means of storing metadata in a machine-readable format. The JSON format consists of key and value pairs (e.g. 'temperature' and '4 degrees') that are arranged within a hierarchical structure.

2. *Contextual metadata formats will likely need version updates*

   As technology continues to develop, additional metrics will likely become more valuable or available for capture to smart devices.

3. *Contextual metadata can be collected by genre*

   Given the abundance of individual data points, establishing 'genres' would allow for similar data sets to be grouped together as a selectable group, reducing a user's need to individually select points to include or exclude depending on a renderer application's needs. This also provides consistent sets of data to developers looking to incorporate a related data set into their applications.

4. *Not all Contextual metadata elements can be captured or stored within an object*

   This is due to the sheer amount of potential elements that could be included, the standard's 'version', Unfortunately for developers, this means that a particular field of information may not be readily available within the framework (although it could be injected retroactively in some cases). This is a natural limitation of contextual metadata, as the amount of data could theoretically and limitlessly expand in domain, memory size, and time required for capture. Creators of objects with contextual metadata would likely prefer not to wait for the data to be collected, when possible. There is also the matter of privacy concerns with the captured contextual data.

5. *Primary-source contextual data parameters are more valuable than secondary, annotative parameters*

   Any information elements that need to be captured with inputs not readily accessible online are more valuable as contextual metadata and should be prioritized in the establishment of a data set.

6. *Primary-source contextual metadata can be used to help verify the validity of secondary-source contextual metadata*

   Since captured contextual metadata effectively acts as a key for annotative metadata, it can be reused as a means of assisting in verifying the validity of secondary-source metadata should there be concerns regarding its origins or whether it may have been manipulated.

7. *A Multimodal Capturing Application can reduce the dependency of rendering applications on external reference APIs*

   Instead of rendering applications requiring API and AI integrations to make analysis of elements of an image, a Multimodal Capturing Application can do this work in advance, only requiring the end applications to reference the required element from the framework. This also enforces standards and expectations for data formatting.

8. *The size of metadata element values needs to be carefully considered*

   There are multiple considerations that need to be accounted for when determining what elements should be included in a standard based on each element's respective value. If the values have no predefined size or restrictions, they could potentially be padded with an excess of information. Many digital file headers (not just their metadata) include descriptions of the size of its elements to assist software algorithms with decoding their contents without error.

9. *Contextual metadata benefits from a standardized source of data*

   To ensure that contextual metadata is comparable across instances, whenever possible it should contain metadata with comparable, scalable metrics. There is also the concern of biased measurements or statements to consider: "if metadata is a statement about a resource, that begs the question of who is making that statement" (Pomerantz 2015, 102).

10. *Embedding contextual metadata into a single digital object results in the creation of a single, easily transferable unit of representative data.*

    This reinforces and makes obvious the relationship between the digital object and its metadata.

I would also like to draw attention to five guiding principles outlined by Marcia Lei Zeng and Jian Qin regarding digital metadata standards and upkeep ([2000] 2016):

1. *Modularity*: allow the protocol to be built in blocks defined by genre
2. *Extensibility*: allow the linking one metadata set to others to provide more detailed descriptions
3. *Refinement*: strive for precision of detail in description
4. *Multilingualism*: to support cultural diversity

5. *Interoperability*: allowance for multiple systems, soft/hard-wares, data structures, interfaces to utilize and exchange the data with minimal loss of content or functionality

A quality metadata standard ideally is in keeping with all of Zeng and Qin's components, and they are presented here to reinforce the value of the principles they represent. While it is not within the objectives of this thesis to accomplish all five of these goals, I would imagine a fully realized contextual metadata standard would encompass these facets and notions to ensure that the resulting framework is multi-compatible with users and systems and is well-structured, organized, and facilitates further development.

Taking these considerations into account, I have created my own contextual metadata framework and format called 'multimodal frames' that embeds external context into a digital image to further investigate these notions through research.

4.2 Creating Multimodal Frames using The Contextual Metadata Framework

Multimodal Frames are the composition of multiple modes of perception into a single file. The format I will be employing in my research takes several liberties compared to the generalized model, with a focus on modifying the digital photographic process on smartphone devices to allow additional contextual metadata to be injected into a single PNG file. This is accomplished via steganography, the process of embedding hidden information into digital files. Steganography's selection as the encoding practice was chosen for the practical purpose of needing the resulting images to be uploadable to social networks without data loss; most embedded metadata is removed upon upload to the platforms and convincing these network giants to adopt a new file format within my timeframe is completely unreasonable.

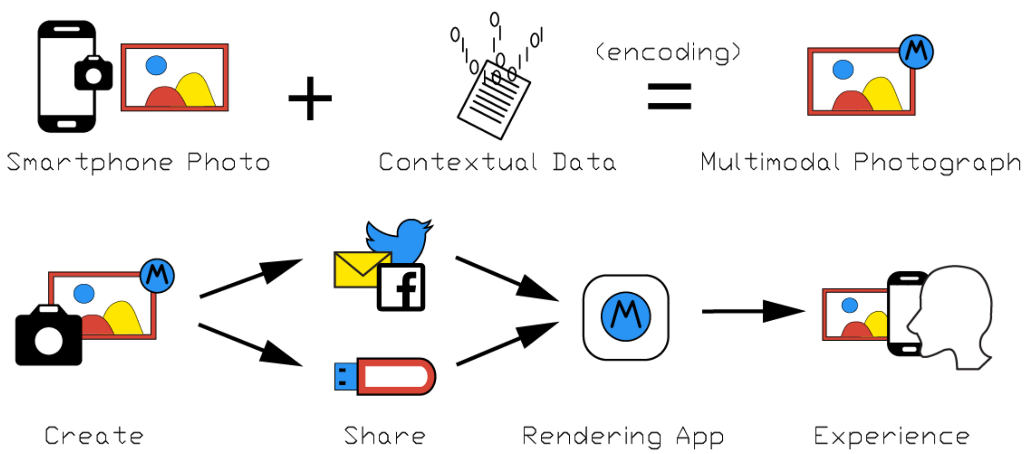


Figure 6: The multimodal media pipeline

The figure above demonstrates the multimodal framework in the form of a flow diagram: a typical smartphone image is captured and combined with related contextual metadata. The contextual metadata is encoded into the smartphone photo via steganography, resulting in a multimodal frame. The second row shows how a newly captured multimodal frame may then be stored or exchanged, then rendered into an engageable experience for end-users of rendering applications.

The amount of data elements that I will be including in my format is also limited in comparison to the numerous elements that I have suggested including into the standard. My goals in the prototyping phase are to demonstrate a select few elements being used effectively for each project's objectives... I imagine that any renderer application would likely not use every data point in such a standard. I will also not be incorporating machine vision for scope considerations. While it is possible to capture and embed social or qualitative facets of cultural history, I am primarily interested in what can be done with concrete values with minimal cultural connotations (e.g. weather, time, geospatial location, pollutants in the atmosphere), aspects of a moment that are comparatively much easier to acquire and quantify given existing technologies.

4.2.1 Steganography and Steganographic Metadata

Commonly associated with espionage and cryptography, steganography is the practice of embedding hidden data into digital files, and for the purposes of my prototyping it is a suitable workaround for the metadata removal trend on social media networks. For the sake of researching this metadata framework, it is important that it is capable of being transmitted on existing infrastructure, hence the necessity to circumvent existing metadata policies. By using steganography, it is possible to hide metadata within files in a manner that cannot be manipulated or erased without compromising the quality of the original media element.

Digital images can be classified as raster images (pixel-by-pixel representations, e.g. JPEG[34]), vector images, or as a collection of raw data of visual information (e.g. the 'raw' photo format). Predominant in photography, the JPEG image is a compressed, 'lossy' approximation of a capture largely based on mathematical cosine formulas and is often significantly smaller in file size than other bitmap images. Another popular alternative on the web is the PNG[35], a non-lossy format with several convenient qualities for steganographic encoding; it is compatible with a wide range of platforms and, compared to lossy formats, is often not compressed or altered by software when transmitted online. Users have an expectation that the pixel data in PNG images will not be

[34] Joint Photographic Experts Group, the name of the group that created the format in 1992

[35] Portable Network Graphics; a format of bitmap image with a channel for opacity frequently used on the Web. In contrast to JPEG, PNG images do not use approximations to determine its pixel representations, instead representing each pixel distinctively within a grid.

compromised upon upload to their platform of choice, which benefits the practice of *lowest-bit steganography* nicely.

The Lowest-Bit technique is one of many popular steganography methods for injecting data into PNG and other bitmap images through the modification of their bits[36]. Each colour in a pixel is represented by the binary representation of a number (00101010 = 42); the idea behind lowest-bit steganography is that the least significant digits (e.g. the last two binary numbers in an 8-bit binary representation) can be modified without significantly changing the appearance of the image. This amounts to approximately a 1% difference in colour value and is applicable to all four channels of colour in the image (red, green, blue, and alpha). Given a 1080x1920 pixel image with four, 8-bit channels and an average of 4.7 characters per word, you can embed a single ASCII character for every pixel, or approximately 440,000 words into this image[37].

---

[36] A bit is a binary digit; either a 1 or a 0. Eight bits make a byte, and each colour channel in an 8-bit PNG image has eight bits; eight bits can represent any decimal number between 0 and 255.

[37] A smaller example: given a 640x480 pixel image and using only the alpha channel, we can get 76,800 ASCII characters or approximately 16,340 words by this logic, which is still a significant amount of information.

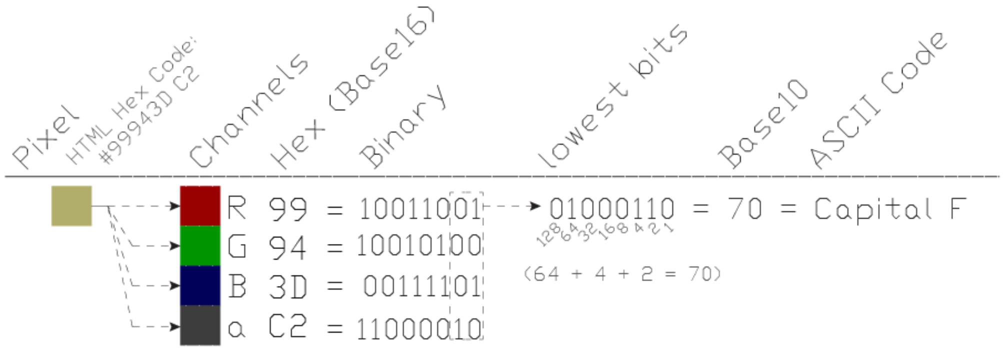


Figure 7: Lowest-bit Steganography

Figure 8: Encoding "France"

This technique of embedding information has been used for all sorts of purposes, including games. The Maxis title *Spore* (2008) is, perhaps, the most famous case. The creators of Spore, a Darwinian-themed simulation game, wanted to give players the ability to create their own creatures and share them on social media. To simplify the process of sharing, the code to generate the creatures was embedded into renderings of the creatures themselves in PNG format with lowest-bit steganography (Batchelder 2008). These images could then be uploaded to social media networks, where they could be freely downloaded and imported back into another client's copy of Spore. The technique has also been implemented in the PC version of *Monaco: What's Yours is Mine* (Pocketwatch Games, 2013), where the level editor can output images of the levels from an elevation perspective, embedded with the coded instructions to import the level into another copy of the software.

5.0 Use Case Discussion

A contextual metadata framework can already be integrated into existing digital capture technologies, as will demonstrated by the Proof-of-Concept prototype in Section 6.2. What remains to be seen is whether (or when) this concept will be integrated into existing technology ubiquitously, which is necessary to properly evaluate whether ECM can really be useful as the contextual metadata framework proposes. As such, we proceed with a speculative investigation of an 'alternate present' (Auger's term, Section 3.2) where such a technology is ubiquitous and fully integrated into existing related technologies[38]. For the sake of discussing how the metadata framework might be integrated into society as well as its benefits and shortcomings, setting the framework against the familiar background of our present day allows for comparisons to similar topics and grounded speculations.

We will take the following assumptions about this alternate present:

- The multimodal framework is ubiquitously integrated into the output image of smartphone cameras for much of the population. Each photo has data embedded into it describing details particular to the moment and location in which it was captured.

---

[38] If we use the Larson & Larson terminology (Section 3.2.2), we can consider this 'the system' for our use cases.

- The types of data to be utilized in the framework are not specific, but we will assume that every capture incorporates GPS location, time, and other data elements that are already incorporated into digital photography.
- Users may opt-out of embedding information into their images if they choose to do so though the settings of their device (although the way this option is provided is left open to speculation).
- The multimodal framework can be transmitted securely on major social networks if it is not deliberately tampered with by external parties.
- The data embedded into these images will be metric, and data collected external from the device will be limited to legal, publicly accessible sources (see *Mobile Metric Data*, Section 4.1.2)
- The multimodal framework is not assumed to be built upon steganography. To reiterate, steganography was chosen as a means of encoding contextual metadata into my prototype’s multimodal frames (Section 4.2.1) because I lack the capacity as a student to enact changes to the existing metadata infrastructure of the web for the sake of research. Multimodal frames could just as well be created based on existing metadata formats or unusual formats that I have not covered in this thesis.

Hopefully this will account for many scenarios that could arise and keep discussion relevant to the material introduced in the context / literature review.

5.1 Overview of User Groups

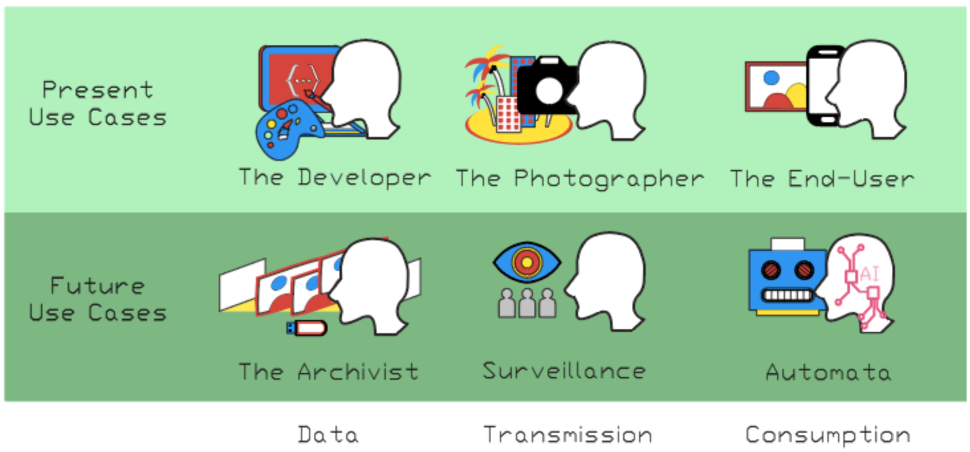


Figure 9: Selected user groups involved in the use case analysis

For the purpose of use case analysis, (see Section 3.2.2), I have separated the user groups involved in a multimodal frame's lifetime and analyze how each role may be affected within the alternate present framed in the opening paragraphs of this chapter. These roles are not intended to be all-encompassing, but rather touch upon several major areas that are expected to be affected by this alternate present. For convenience, each of the user group profiles are accompanied by brief bios that outline the benefits and concerns of their position relative to the contextual metadata framework.

The 'Present Models' include The Photographer, The Application Developer, and The End-User. *The Photographer* represents users of devices capable of capturing multimodal frames during the act of photographing the image. How the resulting artifact is used thereafter by an individual is the discussion for *The End-User*, who may or may not be *The Photographer* as well. The software and rendering applications that these two roles will interact with for capturing, viewing, and manipulating multimodal frames is the domain of *The Application Developer*, a role consisting of developers, designers, and artists. **This is the role that I have embodied in the prototyping section as a means of evaluating the value of the prototype for this particular user group***.* These first three user groups provide a background to the creation and use of renderer applications such as the prototypes and share the most direct relationship with the prototyping process.

Moving on to the 'Future Models': groups and individuals invested in the use of this new media for corporate, governmental, or surveillance purposes are discussed in *Surveillance Networks*, with some overlap with *The Archivist*, a role concerned with data histories, databases, and storage. Finally, we have *Automata*, a role briefly discussing unmanned use of the multimodal framework. These roles are more indirectly associated with multimodal frames, but represent critical topics concerning the future of the proposed metadata framework.

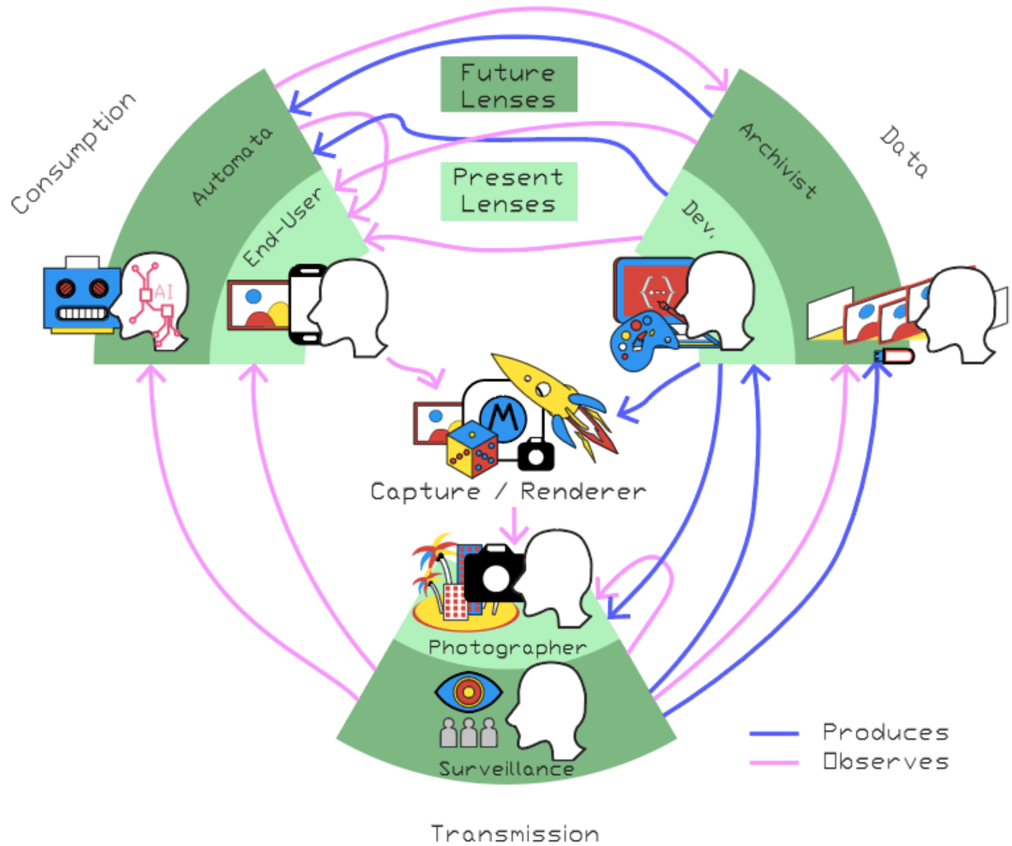


Figure 10: Relationship diagram of the six user groups and the creative prototype

The above diagram roughly describes several relationships that each user group has with respect to one another. I have also organized three categories of user groups that bear similarities in their function:

- Data: pertains to collection and manipulation of data
- Transmission: translating information from one space to another

- Consumption: users who consume multimodal products

There are many overlapping interests shared between each user group. For example, the Surveillance group is interested in monitoring behaviours and history, as well as sponsoring the development and investigation of multimodal application development and study.

5.2 The Application Developer

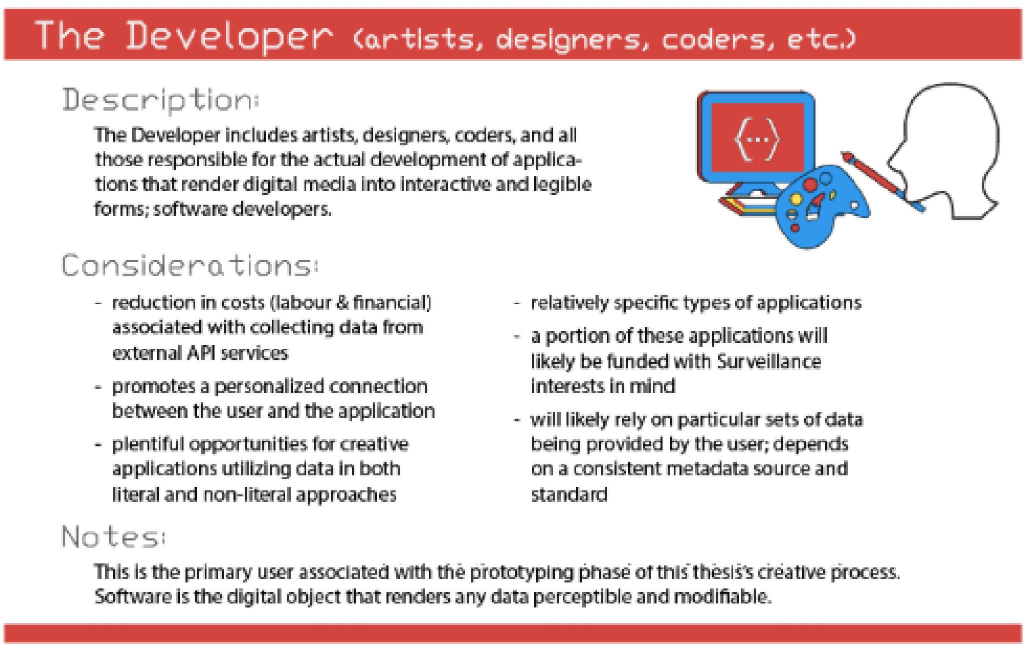


Figure 11: Bio Card: The Developer

As I am effectively embodying this role in the prototyping process, I advise the reader to review Chapter 6 for insights from the process of Research-Creation;

this user group is of primary interest and represents the target audience for the purposes of my research.

Application developers are the designers, artists and programmers[39] responsible for the creation of applications that render digital media into interactive and legible forms for use by End-Users (Section 6.4)[40]. Developers incorporating multimodal data will need to weigh the needs of the application against the degree of literalness required to effectively use the application (see Rendering Applications, Section 4.1.1): both literal and figurative methods of representation have their merits depending on their intended purpose. Without some level of accuracy, *Google Earth* would fail to be an accurate representation of the globe, and without some level of context or relationship to reality it can be difficult to determine the meaning and intention of a renderer and relate it back to external knowledge and experiences, and therefore impact how effectively a person can embody the interface. The multimodal frame caters to both interests by providing contextually relevant data open to both formats of interpretation as a user-friendly transferable unit of packaged data, which contributes to the

[39] I have chosen to lump these two occupations together as the process of developing an application is cumulatively an act of design and requires creative representation techniques developed both through traditional artistic approaches as well as through creative coding. The resulting application should in an ideal scenario merry code and art (are techniques really different from algorithms?)

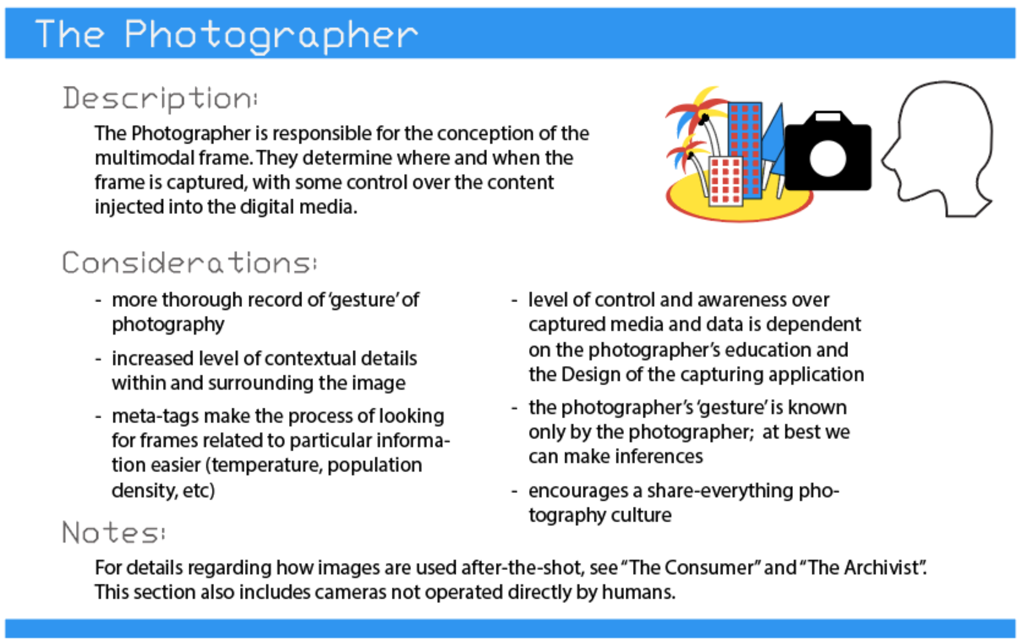
[40] Larger business solutions will be covered in Surveillance Networks, Section 6.6.

development of applications with social interests. By nature of the contextual data being correlated with the position and time in which it is created, it also encourages users to experiment with different approaches to data capture (varying locations and time to produce different results and comparisons in the rendering application).

Application Developers also may benefit from cost reductions associated with the 'single-unit' format of the multimodal image, compared to data collection associated with API calls and proprietary reference libraries (e.g. CoordinateSharp, see Section 8.3). Multimodal media can also promote a personalized connection between the applications and the contributing End-Users, who would likely supply the input data. However, this relationship may also create a data dependency; if the data must be supplied by users, the application's performance related to multimodal media is dependent on the user's active participation.

Ultimately, it is in the hands of the Application Developer to make design decisions regarding how the data will be represented in a rendering application to their or their employer's criteria for creativity, accuracy, affect, and contextual relevance. The allure of these renderers combined with a user's knowledge and tolerance of the data incorporated in multimodal media and the capture application's data privacy management features will likely be the deciding factors for user acquisition.

5.3 The Photographer



<u>Figure 12:</u> Bio Card: The Photographer

Parallel to the practice of traditional photography, Photographers are the creators of multimodal frames. They determine where and when the frame is captured, with some level of control over the content injected into the digital media. The act of capturing a photograph is the act of preserving visual information and, similarly, the act of capturing a multimodal frame is the act of preserving the contextual information relative to a location and time. Since "the objectivity of an image (an idea) can only ever be the result of manipulation (observation) of one situation or another" (Flusser [1991] 2014, 83), we can assert that the photographer holds authorship over the photos or frames that they create. Regardless of the photographer's intentions or relationship with the image that they capture, we are primarily interested to know what the importance of the multimodal framework is to the act of capturing a frame and how this benefits the photographer and their practice.

There exists a difficulty in understanding the contextual qualities of a photograph second-hand. As Terry Barrett describes, "it is difficult for viewers to arrive at a trustworthy interpretation if they don't have some prior knowledge of the photograph: who made it, when, where, how, and for what purpose" (1999, 106). By capturing additional details about its context, multimodal media may, in conjunction with a cleverly designed rendering application, provide additional points of inference into the photographer's gesture and the captured scene itself, thereby improving our understanding of the image, intention, and moment. This benefits the *informed*[41] photographer by providing an increased 'richness' in raw captured data and can potentially display their intentions and the qualities of the scene they are capturing with heightened accuracy. This data

[41] It is arguable whether there is a perceived benefit for a photographer who is unaware of the metadata captured parallel to the photograph, which may be the case depending on how transparently a camera application is developed.

can also be used to assist the photographer in sorting through captured media for contextual qualities (see *figure 1*). However, the increase in the available datasets and rendering applications associated with multimodal media could promote (or lead to protests of) a share-everything, always-connected[42] photography culture, as presently promoted by social photography applications such as Instagram and Snapchat.

5.4 The End-User

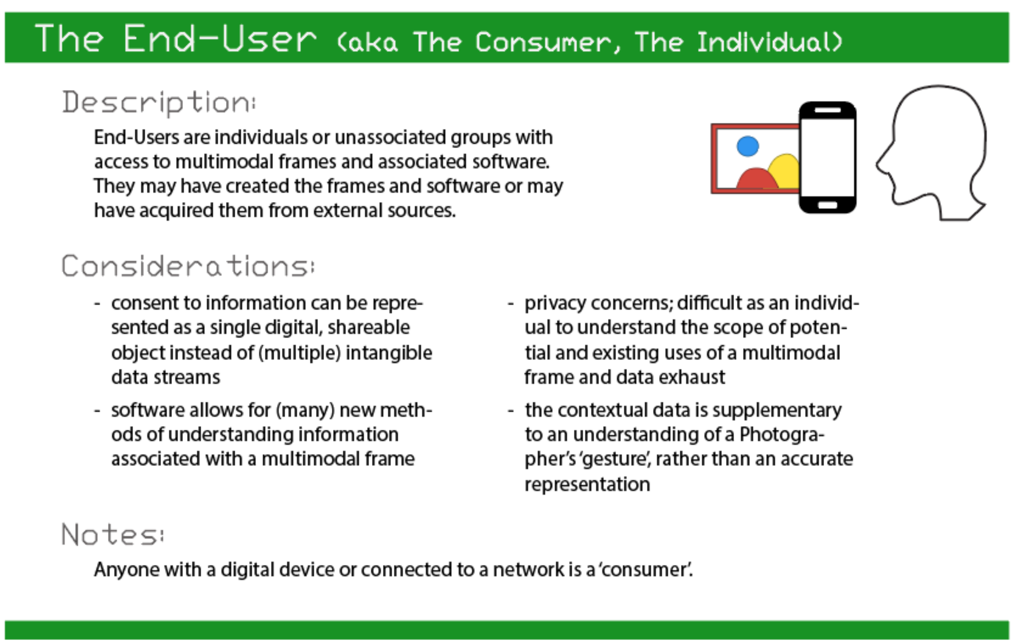


Figure 13: Bio Card: The End-User

[42] "Mobile technologies promote an 'always on' lifestyle, that is, continuously connected." (Arvidsson and Delfanti 2019, 48)

End-Users are a highly generalized user group of individuals or unassociated groups with access to multimodal frames and associated software. The broad scope of this lens makes it difficult to establish a concrete understanding of the individuals it represents, but we can establish several key features and observations that define this user group: people who directly engage with consumer software designed to be compatible with multimodal frames. Paralleling the existing practices related to the storage and movement of digital images (including photographs) can provide us insight into how a multimodal frame might be handled in practice, since their usages largely overlap.

Many smartphone users have amassed a modest collection of images through the process of photography, media creation, or downloading. Some of this data is personal, other elements of this data have been externally sourced, with both cloud- and hardware-based storage mediums. Many of these devices have several renderers installed, allowing users to review these images and their data for informative, aesthetic, and entertainment. The images may act as memories or mementos of past experiences, for communicating socially over networks, for purposes as mundane as an image of a chess set prepared for an eBay auction. There are innumerable uses for data, many of which have or will be outlined in other use cases: surveillance and privacy, promotion and artistry, archiving and socialization, expressions… an end-user could potentially have access to all of these technologies as well as the full spectrum of their boons and concerns. Conversely, this vastness in applications and potential use-cases surely results in naïvety on part of the consumer, or really any individual connected within massively networked technologies in terms of the media they both consume and produce as a trade off for the wealth of knowledge and power they are provided. But being in the public sphere, perhaps the multimodal frame could contribute to a greater public awareness of just how much data can be accessed with a single GPS coordinate in its presentation as a singular, human-readable file unit.

5.5 The Archivist

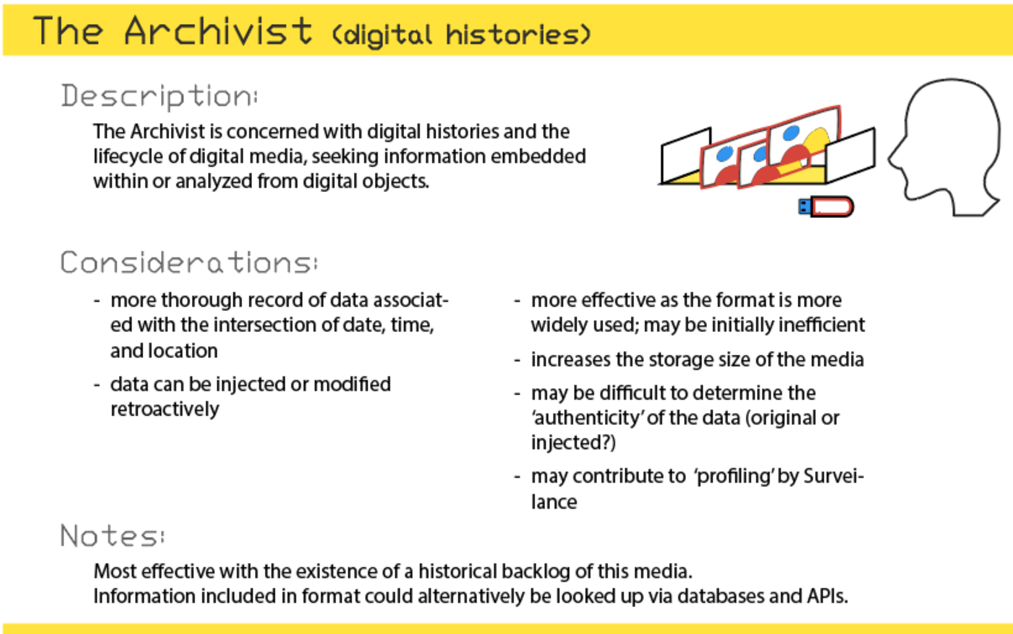


Figure 14: Bio Card: The Archivist

Supposing that the multimodal frame is as ubiquitous as the photograph today, the storage of frames would constitute an enormous and highly networked amount of data[43]. In 2017, Eric Perret estimated that 1.2 trillion digital photos would be taken and 4 trillion digital photos will be moved to some sort of storage medium (Perret 2016). Corporations are constantly creating backups and individuals are increasingly cluttering their hard drives and cloud storage with

[43] Notwithstanding the amount of padding that additional metadata would add to a file's average size; given my uncertainty on the limits of this extra data, it is difficult to ascertain the increase in storage size per photograph

files. Furthermore, websites and the data they store are frequently modified or deleted as domains optimize their content or change ownership[44], and the process of transferring data between platforms is complicated by varying platform-respective methods of managing metadata. Being modular and unit-based by nature, the use of multimodal frames as storage elements may improve the process of transferring relevant information between platforms[45].

In the conference proceedings of *Understanding the Cultural Value of Memory*, Tapio Mäkelä contextualizes this phenomenon by calling the World Wide Web the "largest graveyard of digital and archival desires," noting there is "always someone, somebody who searches in the archives and brings up a name, a person, a location, a thought, or an event that did not exist for 50 years" (Mäkelä 2011, 145). These vernacular images frequently retain some amount of value to their producers: "to the people who took them, they were deeply valuable; to anyone else, mostly worthless. Their peculiar sort of pricelessness made archivists of regular people" (Herrman 2018). Yet, in discussing the practice of archaeology, Mark Shepard remarks that "common objects drawn from daily life do not simply (passively) reflect cultural forces (trends in taste and

[44] E.g. if an image hosting service was bought out by a larger company (Herrman 2018), could no longer afford to cover the operating costs of their servers, or if the data was compressed for storage concerns (Russell 2017)
[45] Zeng and Qin refer to this phenomenon as *interoperability*; see Section 4.1.3

fashion, for example) but also actively participate in shaping the evolving social and spatial relations between people and their environment" (2009, 5). By providing future archaeologists of digital histories additional data to collect, we might enrich the narrative of our historical development. On the other hand, an ever-growing backlog of historical data of living entities also contributes to profiling practices that could fuel a digital surveillance-based police state.

5.6 Surveillance Networks: Corporate and Government Use

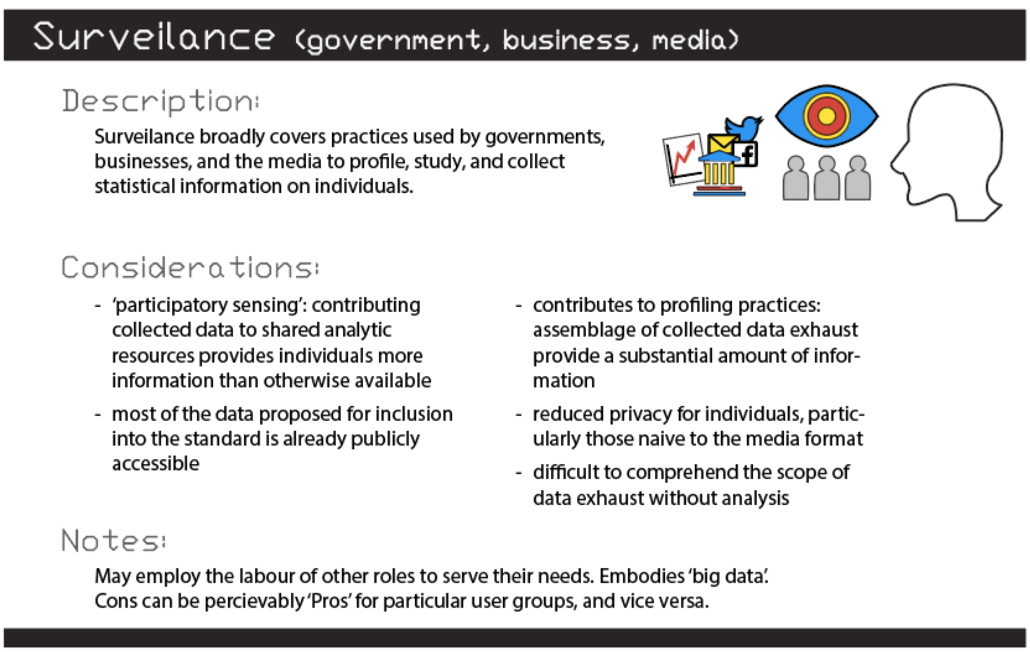


Figure 15: Bio Card: Surveillance Networks

Surveillance broadly covers practices used by governments, businesses, and the media to profile, study, and collect statistical information on individuals. This use case largely consists of entities monitoring networks of information.

The practices of analyzing, processing, collecting and interpreting data from various sources is highly marketable (e.g. BlueKai) and depends heavily on global corporations for client acquisition, habit monitoring, and social behaviour tracking. In the words of Adam Arvidsson and Alessandro Delfanti, "the ability to produce, manipulate, and distribute information becomes one of the main factors related to wealth and power, and therefore a battleground of economic, social, and political development for individuals, businesses, and governments" (2019, 21). As networked individuals, we are participants in these surveillance practices. By being connected with a smartphone, our actions are not only monitored, but can contribute and benefit from 'participatory sensing' practices that aggregate crowd-sourced data to manifest collective representations and visualizations such as the traffic-monitoring system in Google Maps (Yau 2013, 246–7). In exchange for the data necessary to provide useful location services, Facebook users consent to the use of information for other unspecified purposes related to usage statistics and advertising (Facebook, Inc. 2018).

“We kill people based on metadata.” These are the words of General Michael Hayden speaking at Johns Hopkins University in April 2014, the former director of both the National Security Agency (NSA) and the Central Intelligence Agency (CIA) (Pomerantz 2015, 124). Particularly since Snowden’s disclosure of classified documents in 2013, the American government’s surveillance practices have become more transparent; the NSA actively collects data about phone calls directly from phone carriers, including caller & recipient numbers, time and duration of calls, and the locations of caller and recipient (Pomerantz 2015, 124). If a government agency identifies a ‘person of interest’, the metadata about a particular phone can be queried in a phone carrier’s database (usually following a court order) (Pomerantz 2015, 124; Arthur 2011).

When shared and exchanged within a surveilled network, multimodal frames become another unit of data that can be drawn upon to as another datapoint for analysis or profiling. This information can be complemented with existing available metadata to enrich a profile on a person: "dataveillance[46] operations do not require a centralized system, provided a set of different databases are networked and provided that they share the same means of establishing individual identification, so that a single unit (an individual or

[46] dataveillance: "the disciplinary and control practice of monitoring, aggregating, and sorting data" (Raley 2013, 124); a portmanteau of data and surveillance.

number) can be identified consistently across a range of data sets with a primary key" (Raley 2013, 124). Reflecting on my prototype's implementation of multimodal frames, the secondary 'annotative' data elements represent information that can be easily accumulated for dataveillance while the primary 'sensor' data elements act as the keys for collecting this annotative data. Ultimately, everything that is embedded in my multimodal frames can be just as easily collected from the modern smartphone photograph, which suggests that rather than presenting a threat to users privacy, a multimodal frame and its renderers could be revealing of how much information smartphone users surrender to dataveillance practices by sharing their digital photographs online.

5.7 Automata

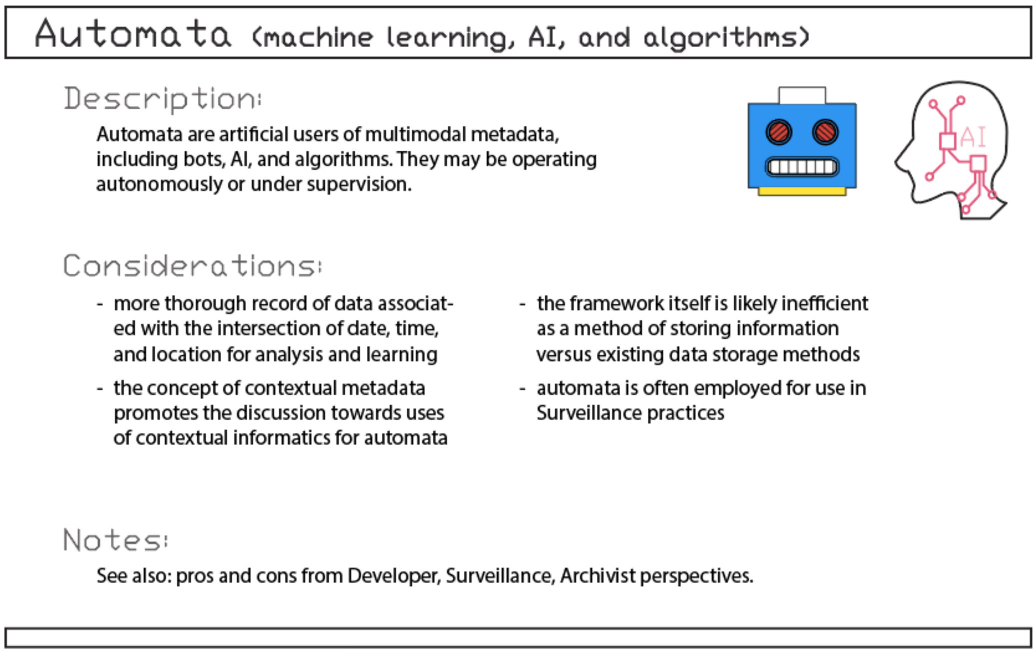


Figure 16: Bio Card: Automata

Automata concerns a group of artificial users of multimodal metadata through automated processes, where an agent utilizing a machine learning process incorporates the multimodal framework as into their learning algorithm. The term Automata refers to relatively self-operating mechanisms[47] ("Automaton" n.d.), which can include agents, artificial intelligence, algorithms, and mechanical processes. Agents are entities that are capable of obtaining a set of percepts or inputs from their environment (real or virtual) using a set of sensors and possess

[47] Contrary to Merriam-Webster's second definition, I argue that their instructions are not necessarily predefined

the ability to act upon their environment using actuators[48] (Castaño 2018, 92). With sensors and actuators, agents have the ability to 'perceive' and interact with their environment without direct human or agent intervention towards the completion of any internally programmed objectives or decision-making processes they may have. Some agents are considered 'purely reactive' and operate without working memory or a history of perceptions (Castaño 2018, 95). In an instance of surveillance, the agent could take capture images or a video stream throughout the day, monitoring simultaneous data and demarqueing its timeline as a series of frames. A comparable study recently used an agent in a farming application: "to detect plant disease and weeds, content-based image retrieval is used to compare with the healthy or useful plant images respectively" (Bhanu, Bhaskara Reddy, and Hanumanthappa 2019).

A multimodal frame could contribute towards a timeline of an agent's data-driven observations, providing an easily transferable, packaged format of information storage. However, following the multimodal framework likely does little to improve over existing frameworks of information transfer and information collection. In fact, storing visual information into human-readable packages is likely inefficient for storage and machine analysis purposes compared to methods that could investigate raw capture data. The concept of

[48] Actuators are based on the verb 'to act'

utilizing additional contextual data, however, is proving useful for remote-operated robots such as urban rescue robots and military drones, where "panoramic visualizations, high-fidelity imagery beamed into heads-up displays, and additional sensor information such as temperature, slope, and ambient noise are all being used to make the operator feel present enough to make good decisions" (Nourbakhsh 2013, 75). The additional datasets could assist in reinforcing learning algorithms and providing content for data scraping processes. As Shepard describes:

> "Real-time [contextual] information [can be] correlated with historical data of someone's mobility patterns, purchasing history, personal interests and preferences (as reflected by user-generated profiles) in order to make more accurate predictions about what his or her wants and needs may currently be, or what actions s/he is likely to take next." (Shepard 2009, 4)

Of course, these Automata could just as well be employed for surveillance practices. Anind Dey and Gregory Abowd state that "a system is context-aware if it uses context to provide relevant information and/or services to the user, where relevancy depends on the user's task" (2000, 6); this definition clearly alludes to how automata could be used for dataveillance purposes. As HostingFacts reports, a staggering 51.8 percent of all Internet traffic comes from bots (the remainder reported as 'human') (2018); considering the number of active internet users worldwide, this is an immense figure.

6.0 Renderer Prototyping

Please refer to Appendix B for an elaborated process journal of my design notes and decisions; the content in this section will be condensed to reflect core observations that are valuable towards the research objectives of this thesis.

This chapter serves as a description of my research-creation process and a summary of my associated findings. The objective with the prototypes is to demonstrate a 'high-fidelity splinter' of the possible manifestations of the multimodal framework. The 'refined' demonstrative prototype is a creative presentation of the research that I have conducted, built to work with present technology while utilizing the contextual metadata framework presented in Chapter 4. This prototype functions as rendering software that effectively converts multimodal frames into playable, procedurally generated game levels, documenting the contextual data of a frame with literal and figurative representation aspects. Effectively, I will be creating a 'bare bones' multimodal camera application as my proof-of-concept and a geolocative mobile game as my demonstrative prototype.

I will be noting my findings through the research-through-creation process through the perspective of an 'Application Developer', one of several users that are covered in my use case analysis, elaborated in the previous Use Case Discussion chapter. This chapter will also discuss the multimodal framework in more generalized terms compared to the niche to be investigated and demonstrated in the prototyping process.

6.1 Design Considerations

Both prototypes (the camera app and the mobile game) consider concepts and ideas presented in a variety of fields that are not directly associated with the core context and literature of this thesis. I consider these to be influences, or design considerations, that have shaped the way I have created my prototypes and their intended audience; this section is dedicated to brief reviews of related fields and connections of interest.

As I will detail in Section 6.3, I will be developing a mobile, geolocative game incorporating a multimodal camera application as a data resource. Users will imaginatively use the camera in tandem with *Vernacular Mobile Photography* practices. The data collected from the produced multimodal frames will then be used to make decisions in the level generation process, accomplished via *Procedural Content Generation*. The use of parameters from reality to drive design decisions draws from the architectural style of

Parametricism. Conversely, the use of spontaneity and unexpected results are more in line with the philosophies behind *Glitch Studies*, where incidental data visualization contributes towards artistic creativity. I will be discussing the italicized subjects in the following subsections to better illustrate the design considerations for the prototype.

6.1.1 Procedural Content Generation (PCG)

Procedural Content Generation is a term for automated processes of creation, driven by an algorithm and associated input data. The term is often associated with the digital games industry and the trend of 'industrializing' the process of level or object creation. PCG strategies can allow developers to create seemingly infinite variations of designs based on a descriptive set of instructions and sequences. This compliments the prototyping interests of providing unique, contextually relevant experiences in software while factoring in multiple modalities of data or experience.

The engineering process of creating PCG algorithms can be considered a 'meta-problem', which challenges the programmers to understand how the content they wish to create functions in isolation and in relation to external factors (as opposed to a more 'hand-made' approach). There are a variety of methods of dealing with the generation of content (Backus 2017; Watkins 2016), but it is usually performed with some aspect of randomness[49] or input data to control the decisions carried out by a generative algorithm (Bucklew 2017).

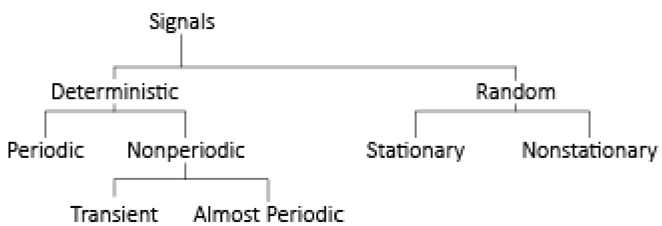


Figure 17: Types of Signal (Hammond and White 2015, 146)

If we look to the signal-processing hierarchy described in *figure 17* and substitute the term 'signal' for input data, we can consider the evaluation of the format, domain and 'behaviour' of this input data as a means of identifying common, patterned, or expected value ranges for consideration in the development process of this PCG algorithm.

[49] it is important to note that 'randomness' is not always completely unpredictable: in some cases, random noise can demonstrate patterns or be restricted to particular domains.

Figure 18: Spelunky HD: Level Map from World 2 (Mossmouth, 2014) (Yu n.d.)

In game development, there are several utilitarian reasons to use PCG: to save time, for expansibility, for replayability, to reuse code, and for rules enforcement. Because of the inherent unpredictability of these inputs, however, designers may opt to hybridize the procedurally generated content with handmade elements (Grey 2017): in David Yu's *Spelunky*, clusters of levels are handmade, then patched together in a process called 'quilting' to create the caverns that players explore (Grey 2017). You can see the results of this process in *figure 18*, where a level is created from a 4x5 grid of 'rooms' with a discrete path from the player start location to the level's exit. A similar process will be used for the purpose of generating levels in the Demonstrative Prototype.

In conjunction with values fed from a pseudorandom number generator, the generated content can be consistently reproduced across various instances of a software provided that the generators utilize the same 'seed'. As Brian Bucklew describes, "pseudorandom number generators start from an initial seed value and then perform a mathematical operation on that seed value that produces the next value in the random series and increments the seed in some way so that the following call will produce a different (but statistically random) value. Whenever you give a pseudorandom number generator a particular starting seed, it will always produce the same sequence of random values" (2017, 272). This allows for a combination of both randomness and reusability in the outputs of a PCG algorithm.

My decision to use PCG rests on the need to provide experiences catered to very specific, innumerable data combinations and the obvious game design benefits for the game aspect of the prototype. Rather than designing these experiences manually, PCG allows me to design writing algorithms to do the design work automatically under highly varying conditions. In practice, a developer utilizing multimodal data would likely run into similar problems of having to design contextually relevant visualizations for such varying conditions; incorporating elements of PCG into the design work is an effective means of automating the workload and allowing for interesting, emergent, and unique properties to appear as a by-product of the process.

6.1.2 Parametricism

Parametricism is an architectural movement that proclaims that design should be driven by data and subject to modularization via variables (Schumacher 2015), and incorporates analysis across a variety of domains (spatial, weather, social, etc.). Patrik Schumacher, the most vocal advocate for the movement, declares that "the opportunities that the information age offers are the new computationally supported information processing, design, engineering and fabrication methodologies that can be brought to bear on architecture's new challenges" and argues that parametricism is the only style that adequately responds to the social needs of the information age and adapts to the vast amount of information that informs architecture (Schumacher 2015). In a sense of the word, parametric design is the architectural parallel to Procedural Content

Generation in its use of data analytics to define all types of elements of a building.

Although this thesis investigation is not inherently architectural, Parametricism's realism, data-driven interest in procedural design, and its relationship with PCG methods are of particular interest as design considerations: PCG discourse largely discusses virtual contexts, whereas Parametricism is largely interested in adjusting to physical phenomena for the sake of creating physically creatable spaces. Since we are dealing with properties of reality as input data for the PCG process, it makes sense to look towards parametric practices for inspiration towards how particular properties of reality are dealt with when making parametric design decisions.

6.1.3 Glitch Studies / Glitch Art

Glitch art is artwork composed of digital errors, whether deliberately invoked or accidental in nature. While rendering applications generally seek some degree of faithfulness to the data they seek to represent, Glitch Studies is interested in corrupting elements of the pipeline between data and representation to create an effect that obstructs their intended purpose. Rosa Menkman, a prominent figure in glitch studies, challenges the ongoing evasion of entropy in signal processing and data management practices by encouraging the exploitation and corruption of all varieties of technological media (Menkman 2011, 11). Because we are interested in investigating how data can be represented by software in a creative manner, it begs the question of how it may be *mis-represented* and still be valuable as a creative medium. Accordingly, the deliberate misuse of data and processing methods can create unexpected or aesthetically interesting artistic works. Of course, my interests in Glitch Studies are not to destroy or damage the framework that I am investigating or the devices/software associated with it, but rather to validate the use elements of the data that I have collected in manners that are unrelated to their context and intended form of representation as a deliberate creative artform.

In a sense, adopting Glitch Studies as a design consideration acknowledges the fact that unexpected, emergent properties resulting from a PCG process can have artistically interesting properties in their *misrepresentation* of reality, and that these properties should be considered valuable among more realistic representations.

6.1.4 Vernacular Mobile Photography

Photographs taken by the average individual have a personal relationship to their photographer. A shopper may take photos of themselves in different dress to comparatively decide which outfit merits a purchase; a professional photographer may capture artistically framed landscapes for the purpose of a gallery showing or as stock material for an advertisement; anyone may take a photo of their to-do list to keep track of the tasks they wish to accomplish. The value of these photographs is also purposeful[50]: the picture of a to-do list might struggle as an artistic statement but is incredibly valuable for its user to keep track of their lives. For me personally, photography provides the motivation to go outside and engage in light physical activity with the quasi-productive context of creating artwork (see Evans 2018, 287).

Relating to professional photography, Batchen describes vernacular photographs as “the popular face of photography, so popular that it has been largely ignored by the critical gaze of respectable history” (Batchen 2001, 57). He is, of course, referring to the fact that these photos that everyday people capture on a casual or personal basis are considered a separate class to the celebrated, gallery worthy images that receive publication: “as a parergon, vernacular

[50] The indeliberate photo has an indeliberate purpose until it is seen. To ascribe purpose is to ascribe meaning.

photography is the absent presence that determines its medium's history and physical identity; it is that thing that decides what proper photography is not" (Batchen 2001, 59). The value of each respective vernacular photo may be debatable, but they collectively contribute to a narrative of human social and cultural history.

Taking a photograph of an experience allows you to take some element of it with you and preserves the frame in time, "miniatures of reality that anyone can make or acquire" (Sontag [1977] 1990, 4). The photograph lacks many of the key elements that would be necessary to fully recreate the moment captured but this *image* of one's perspective of a space remains an authentic captured experience. With the mobile phone, "even the most banal moments of the day can become a point of photographic reverie, potentially shared instantly." (Rubinstein and Sluis 2008, 9); mobile phones have become the largest producer of vernacular photographs, contributing to the social practice of taking photos as memoirs of travel, a medium to share experiences with friends, family and others, an opportunity for reflection upon past events and activities.

Figure 19: Dundas Peak, Dundas, ON, March 2018

As Susan Sontag notes, “taking photographs fills the same need for the cosmopolitans accumulating photograph-trophies of their boat trip up the Albert Nile or their fourteen days in China as it does for lower-middle-class vacationers taking snapshots of the Eiffel Tower or Niagara Falls” ([1977] 1990, 9). The act of tourism, propelled by photographic endeavors, also defines the identity and physical status of famous landmarks, "shaped in great part by the fact that people travel long distances to visit them" (Frith 2015, 86). From personal experience, the parkland of Tews and Websters Falls and the accompanying Dundas Lookout dramatically change with the site’s growing tourism over the years, having grown up in the surrounding neighbourhood. Parking lots were first introduced on-site in the late nineties, then gradually converted to shuttle bus drop-off zones in the 2010s to accommodate the overwhelming vehicular traffic to the park. The foot traffic on the trails themselves have ballooned accordingly, as demonstrated in *figure 19*, though the image contrasts the panoramic views that drive so many people to visit the site on a regular basis[51]. While there are certainly multiple factors behind this phenomenon, the second-most apparent characteristic of these visitors seemed to be their interest in photographing the scene with their smart devices, imaginably to share with others or hold as a keepsake. Vernacular photographic practices like tourism are some of the main drivers for digital photography and encourage the photographer to physically displace themselves (either to capture or to experience), potentially providing access to a broad range of external context that remains undocumented in existing digital cameras.

Vernacular mobile photography represents a parallel to how I would expect my prototypes to be used if they were distributed publicly and successful in their design. These practices are representative of a dominant share of the photography widely exchanged on the internet. I hope to use this understanding of vernacular mobile photography to help imagine how end-user practices may

[51] Contrast the picturesque images with the feedback on AllTrails' website: https://www.alltrails.com/trail/canada/ontario/dundas-peak-trail

impact the use of my prototypes and adjust the design of my prototype and multimodal frames accordingly.

6.2 Proof-of-Concept Prototype

The Process Journal for this prototype is available for reference in Appendix B.1.

In the summer of 2019, I developed a proof-of-concept prototype to demonstrate a working example of a digital camera application that embeds locative[52] contextual metadata into a captured image. Operating on both computers and mobile devices, the information is stored almost invisibly using lowest-bit steganography (see *figure 21*) and can be decoded back into usable contextual data. Evoking notions embodied in parametricism, the application also provides a representation of the captured media in the form of a visual weather simulation as well as a gallery feature to review images previously captured by the application or imported from external sources. Despite the user interface's low-fidelity proto-state, the application is ultimately functional and demonstrates how immediately such a concept could be integrated into existing

[52] As described in Section 4.1.2, Mobile Metric Data, there are many other potential data sources available for use, but for the sake of quickly producing a prototype I chose to utilize a select few, with emphasis on locative data.

capture technology.

<u>Figure 20:</u> Screenshot of the proof-of-concept prototype.

Figure 20 shows the application displaying captured contextual information embedded within a multimodal frame. The photograph was captured in the evening during a cloudy day, which is reflected in the app background. For the purpose of demonstrating the app gallery working effectively for varying times and locations, I embedded data into an existing image I captured at Watkins Glen, NY as displayed in *figure 21*.

Figure 21: A sample of an image embedded with contextual information via steganography.

The proof-of-concept prototype successfully established that it was indeed possible for me to collect and embed ECM as part of a multimodal format into an image on a mobile device and transmit the data across social media networks without data loss. This assured that the framework could advance and be used in a more developed prototype, with a greater freedom to explore rendering possibilities for a multimodal frame. Given the relative simplicity of the task, it also confirmed my suspicions that it would be relatively easy to start including additional data points to existing metadata formats in the sense that this innovation could be implemented in today's technological ecosystem.

6.3 Demonstrative Prototype: Gaia Gate

The Process Journal for this prototype is available for reference in Appendix B.2.

Please note that this prototype extends upon the previous prototype work.

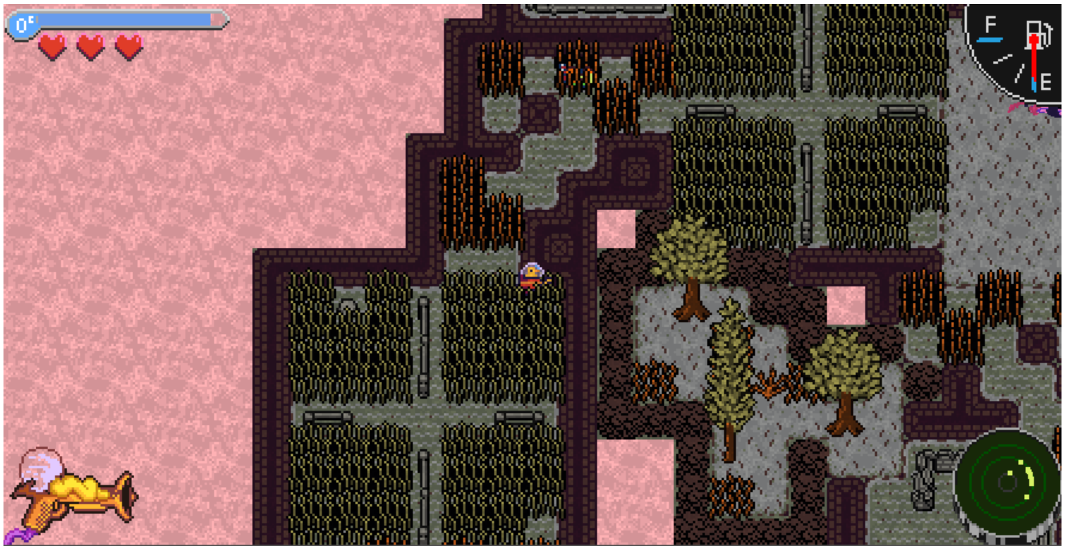


Figure 22: Screenshot of Prototype, featuring a temperate deciduous biome and a croplands anthrome with an overcast sky.

Figure 23: Screenshot of Prototype featuring a grassland biome, with photo

Gaia Gate is an ongoing endeavour to develop an application that utilizes the proof-of-concept prototype's contextual metadata framework towards the creation of a mobile application to display figurative representation(s) of the embedded data in the format of a game. The format of a game was selected to promote a sense of creative freedom in how the data is utilized and emphasize engagement between the user and the media. Of course, this is only one format of many that could be investigated for such outcomes: the project can be considered a 'high fidelity splinter' of the large scope of this thesis investigation, examining one possible product of a speculative scenario where multimodal frames are ubiquitous. As this is a largely solo endeavor, the aesthetic and gameplay scope of the project are intentionally low fidelity to focus on quickly addressing my research goals, which appear to be manifesting in the level design aspect at this time.

The project is intended to act as an exploration of the influence of several contextual data metrics in the development of the rendering application as well as an opportunity to note and investigate any findings or challenges that arise in the process. It is my belief that these findings will reinforce the value of the theoretical framework for metadata described in Chapter 4 for prospective Application Developers interested in contextual expressions of data.

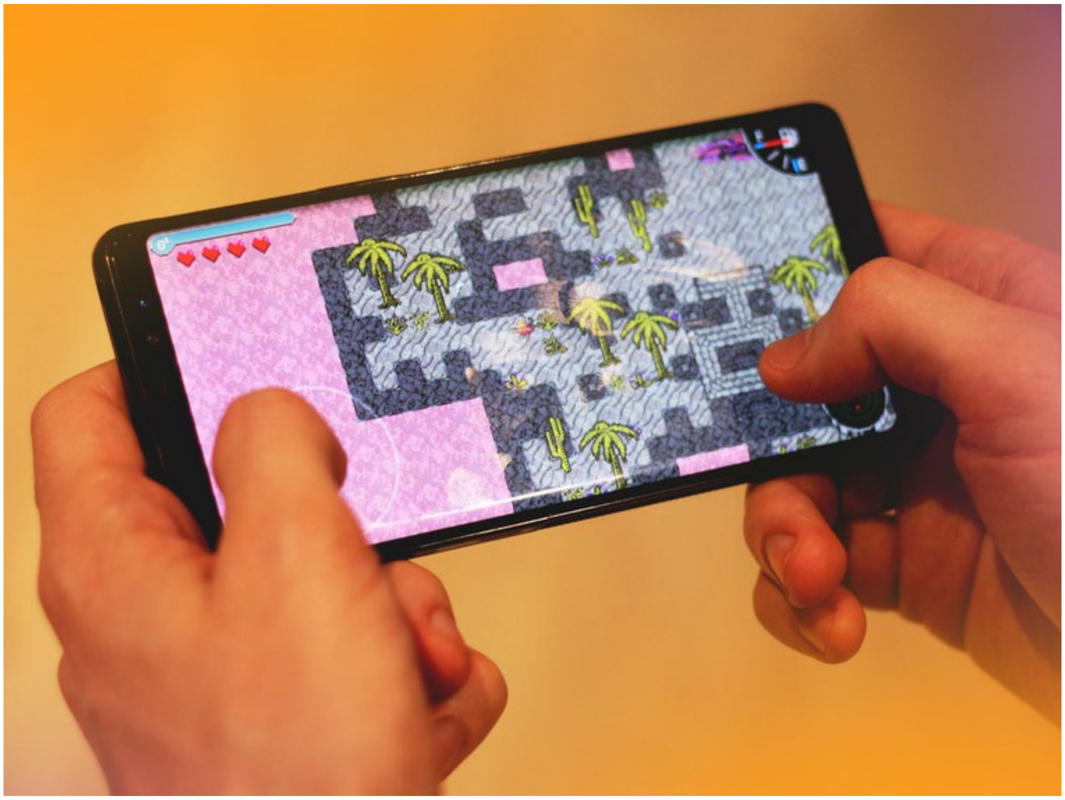


Figure 24: The prototype in play on an Android phone (desert biome)

Given the emphasis on level design to illustrate notions of context, the Gameplay aspect has taken a back seat to the representations of locative data in level generation in keeping with scope considerations. While the prototype is indeed a *game*, the priorities of the prototype for the sake of thesis research lie in the conveyance of the properties of the multimodal frames towards the creation of game levels. Still, I added a simple gameplay element to provide a degree of motivation to explore the levels: using the radar as a guide, collect enough ‘Fuel’ items to fill up the fuel indicator in the UI before exiting the level, before the player’s oxygen runs out.

6.3.1 Prototype Framework & Camera Improvements

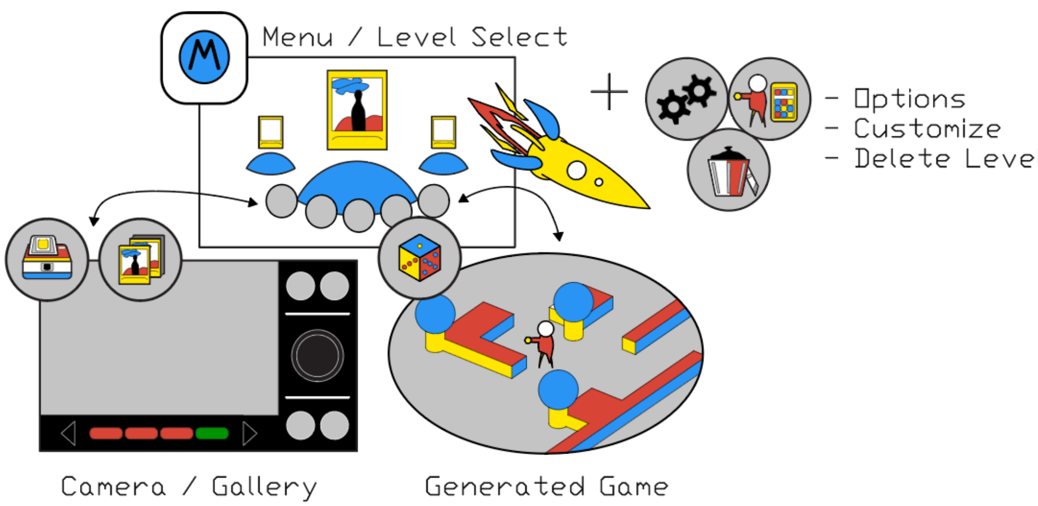


Figure 25: Organization of demonstrative prototype scenes, including the Menu / Level Select, Camera / Gallery, and the Generated Game.

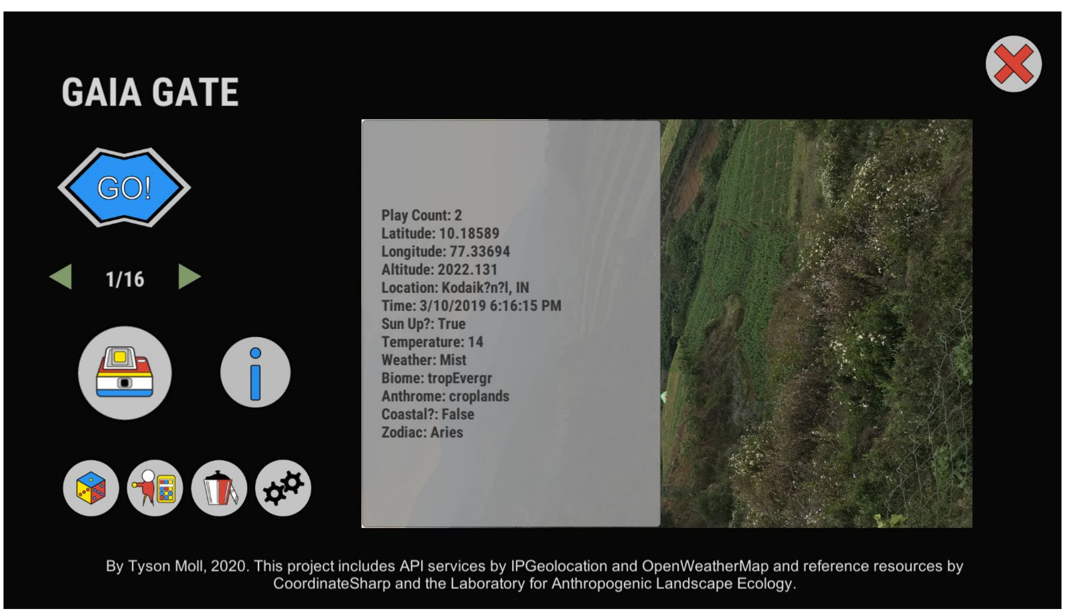


Figure 26: Main Menu of Gaia Gate, with level data created from croplands in Kodaikanal, India (image courtesy of Shikhar Juyal)

The demonstrative prototype is organized into three sections: The Menu, the Camera, and the Generative Gameplay. Upon opening the application, users are initially presented with the Main Menu as depicted in *figure 26* where they may adjust settings such as the level they wish to play (as represented by a multimodal image), adjust various settings, generate a random level, or capture new contextual image-levels using the built-in multimodal camera.

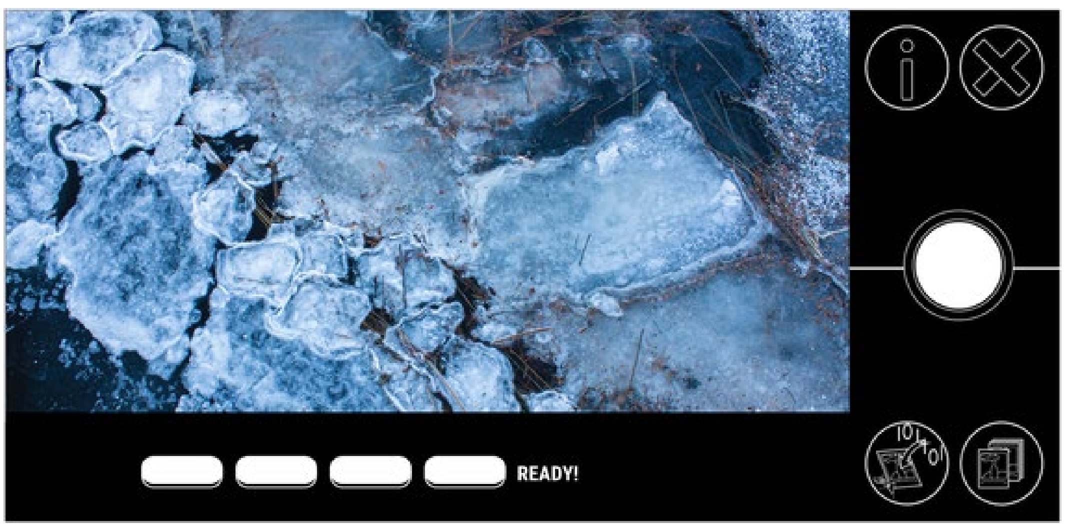


Figure 27: The updated camera mode (dark theme)

Continuing from the work produced in the proof-of-concept prototype, Gaia Gate reuses the camera feature for the initial capture of multimodal frames and contextual data but extracts only the relevant datasets for the purposes of generating levels. This information is stored in a serializable object class for saving binary data and the photographic image is stored as a thumbnail for visual reference. This naturally reduces memory demands when handling the data across the game application, since only the required data values need be saved in the game's memory. This data is then presented as the levels that can be selected in the menu, and the dataset is accessible by the generative algorithm when the game level is initialized.

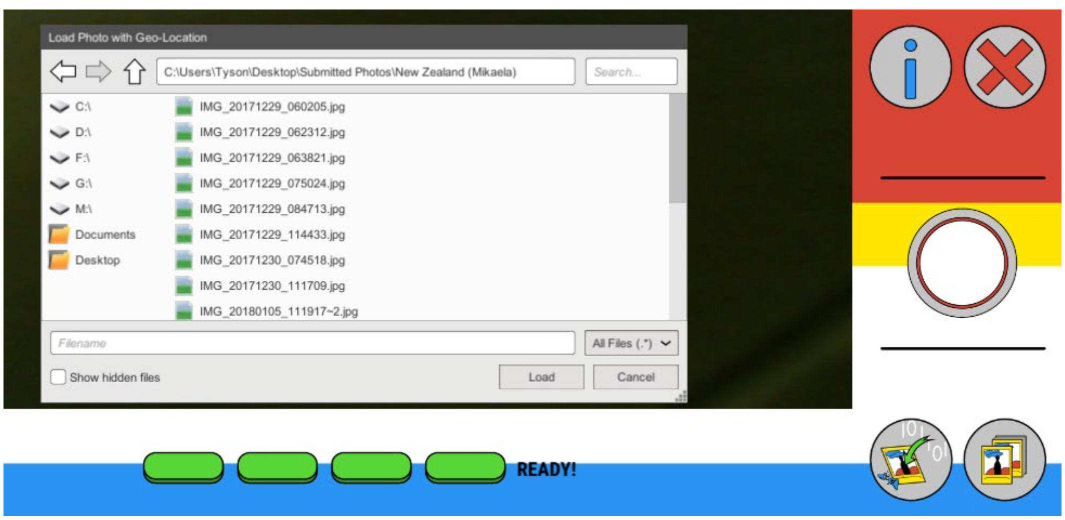


Figure 28: Importing a geotagged image for conversion into a multimodal frame

Figure 29: Gallery mode, with details for an image from Reykjavik, Iceland (image courtesy of Mikaela Manley)

The camera has also received a visual upgrade from the previous version as well as the addition of a photo import feature, where users may import their pre-existing geotagged digital photos into the application to be converted into multimodal frames. It should be noted that historical weather data has not yet been implemented into the prototype due to challenges in finding a suitable source, so the application defaults to the current weather.

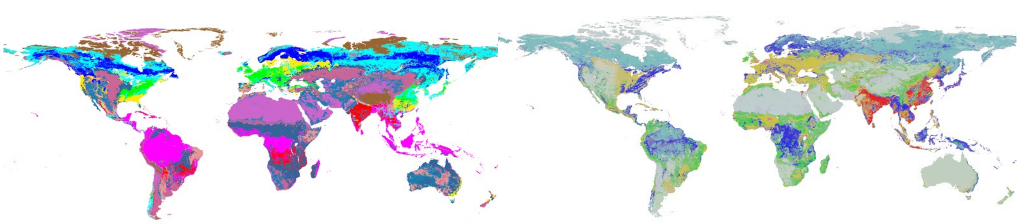


Figure 30: Rasterized versions of ASCII ArcGIS Tables for Biomes (left) and Anthromes (right), With Regions Differentiated in Color (created from the ASCII-format dataset by Ellis et al. 2010)

I have also added the ability to read ASCII- and Raster- formatted ArcGIS tables to the prototype which has increased the number of accessible datasets that I can draw from. In particular, I have chosen to use anthrome and biome mappings created by Ellis et al from a report on “Anthropogenic Transformation of Biomes” (2010), which has enriched the breadth of locative information accessible to the prototype. Lastly, the application now references CoordinateSharp for data parameters related to celestial objects and IpLocation.io for locations based on a desktop / laptop computer’s IP address (solving an issue concerning how computers might access location-based datasets).

6.3.2 Level Design

The core decision-making involved in the prototype with regards to the contextual data captured in the multimodal images occurs in the level design associated with the procedurally generated levels. Multiple aspects of the design of the gameplay space that players participate in is influenced directly by the data they import and respectively how the algorithm for PCG uses these data points to determine how the level is to be developed.

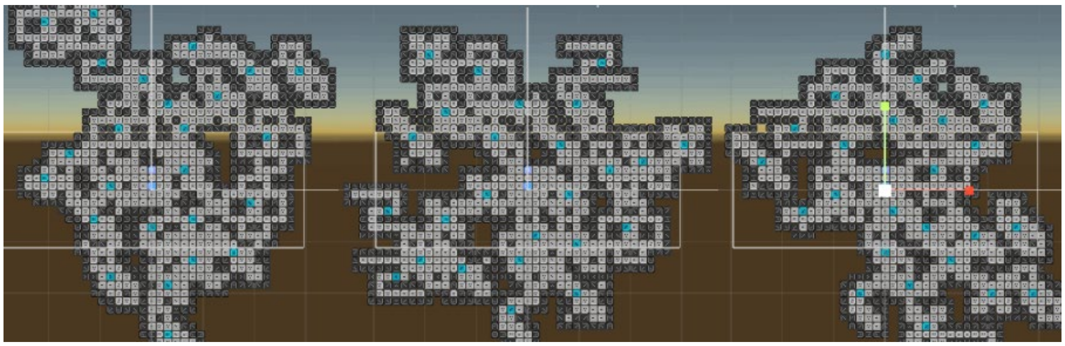


Figure 31: Three early variations of PCG level layouts based on GPS seeds

To ensure that the rendered content is unique between different frames yet maintains consistency between similar values, I have used Procedural Content Generation (PCG) in conjunction with a pseudo-random number seed (see Section 6.1.1). By using a pseudo-random number generator instead of pure randomness, we can consistently get the same sequence of 'random' decisions every time the level is loaded, or any time a multimodal image is loaded with the same set of coordinates. The pseudo-random number seed is made up of the GPS coordinates of the image; by rounding the accuracy to particular digits, we can ensure that every location spaced approximately 100 square meters[53] is generated uniquely by the PCG algorithm. This ensures that fluctuations in position reporting caused by signal interruptions are insignificant (e.g. by urban buildings, canyons, precipitous weather) and ensures that players can recreate common experiences by capturing frames in relatively similar positions on the planet.

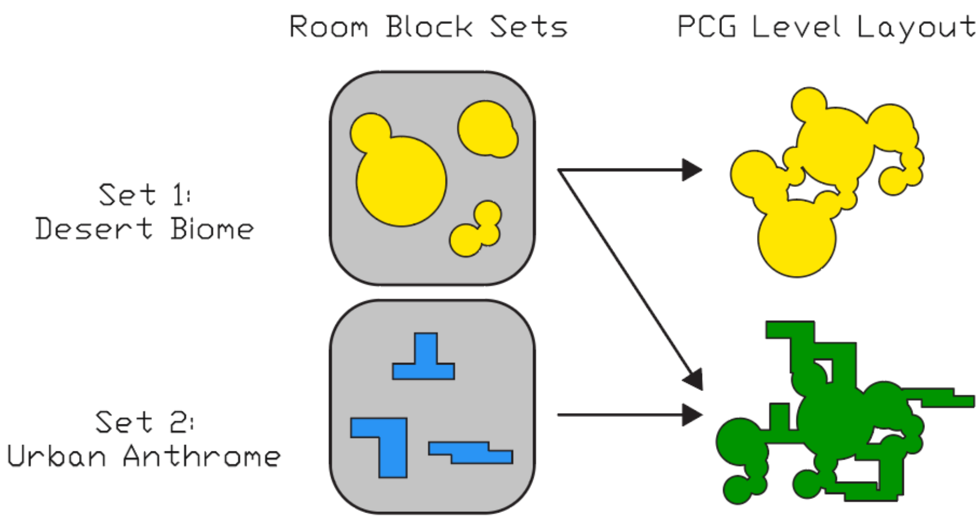


Figure 32: The demonstrative prototype level generation method

The algorithm used for creating the rooms involves blocks representing rooms placed at the world origin and forced to settle out by randomly moving in a cardinal direction, one step at a time, until there is no longer a collision with an existing room block. Once all rooms have settled, the level's walls are generated

[53] For reference, the distance of a degree of latitude is approximately 111km. To achieve a level of accuracy of 111m, an accuracy of 1/1000th of a degree is required.

around the perimeter of the vacant elements of the rooms to fence in the game elements. The room blocks themselves are chosen randomly from handmade template groups for each biome and anthrome, each with their own set of graphical tiles for wall and floor elements. The anthrome determines the ratio of biome rooms to anthrome rooms to be spawned: a city will have a large percentage of its room blocks be anthrome-based, whereas a village will have a larger number of biome-based blocks. Similarly, the biome selects the variety of vegetation objects that will be created in a level.

Figure 33: Early demonstration of biome / anthrome ratio level generation in a Desert biome with a 'smoke' weather condition + level map.

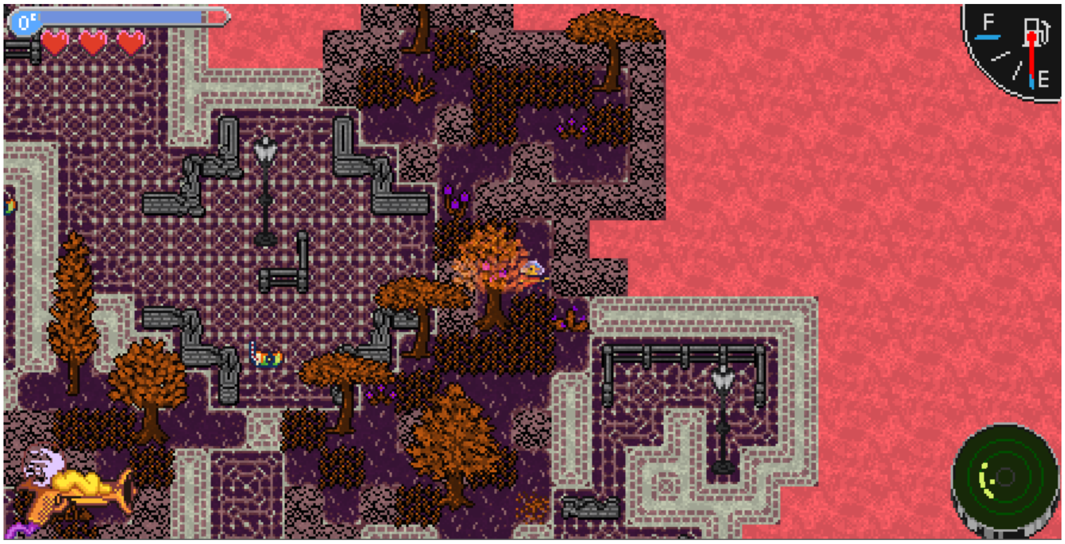


Figure 34: Urban Anthrome in a Temperate Evergreen environment, with overcast weather. Created and captured in Toronto, ON.

Together with the temperature, the biome of a location is also used to determine the colour tone of vegetation. Colder climates will tend towards tones of yellow, brown, or blue, whereas warmer climates will tend towards greener colour schemes. Temperate deciduous biomes have a special colour tone accessible in the fall: reds, to match the colour-changing of the fall. Deciduous trees will also drop their leaves if the weather becomes too cool.

Figure 35: Palm tree and tall grass

The background and colourization of the floor and wall tiles are based on an algorithm I developed several years ago that creates random, visibly distinguishable colour palettes by shifting hue, saturation, and brightness components of a series of colours in calculated measures. I use these to replace a tile's default colour set with my own 'swap' palette sets. As of writing, level tiles for walls and floors are given colour schemes that for the most part do not relate to the data (besides the GPS pseudo-random seed). On one hand, incorporating data related to time such as sunrise and sunset could lead to an interesting relationship between temporal elements and the level's colouration, but on the other hand, the highly varied colour sets means that individual levels within a short locative proximity of one another can be uniquely identifiable and aesthetically interesting. This suggests a sort of conflict between favouring temporal or locative aspects for colouration decisions as a means of emphasis:

there needs to be a creative solution that has some sort of balance between these factors without compromising the unique appearance of various locations. My current solution is to have Day and Night control how brightness is applied to the floor and wall palette generation and the ambient background colour but maintain the randomness of colours being associated with GPS position: reversing the application of brightness to the individual colours results in a modified *feeling* of light within the level.

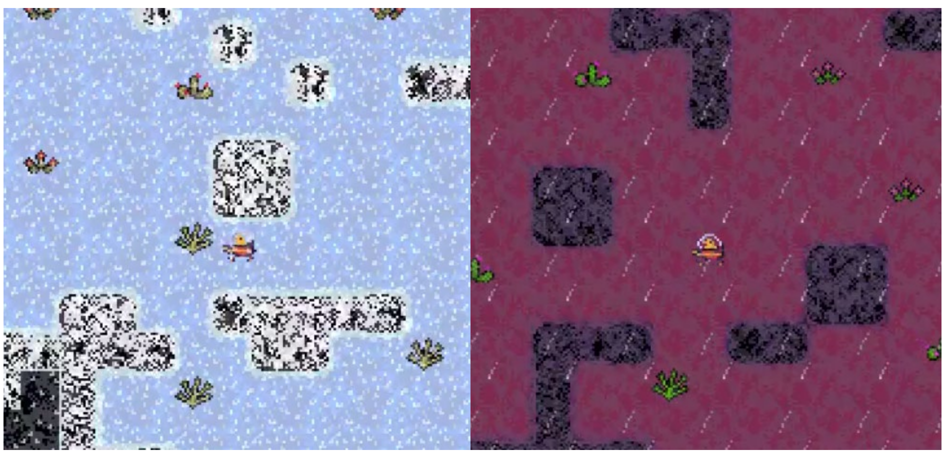


Figure 36: Snow and rain weather conditions

Regarding the weather: if a weather condition is present, it is represented in-game with appropriate visuals. I have plans to modify the direction and intensity of the visuals based on factors such as precipitation, wind speed and direction, and colour cover, but they serve to accurately represent the weather condition of the multimodal image for the time being. From a gameplay perspective, they apply an interesting visual dynamic to levels by obscuring elements of the game and could possibly affect the generation of hazards in the future.

The last element related to level design I would like to mention is related to 'loot', such as coins. This is yet to be implemented, but I currently have access to an ArcGIS table that represents the income levels as globally mapped across the globe. If such a dataset were to be integrated with the generation of said loot, this could create virtual worlds where some regions (and by association, players) are disproportionately wealthier than others. I imagine this could manifest into a controversial situation where developers could make their systems proportional to local welfare to highlight or criticize real world conditions or inversely proportional to balance access to real-world resources (such as ability to travel or finances) with virtual currency. This latter case seems especially ripe for speculation: for example, it could incentivize tourism to welfare states or facilitate proportional siphoning of microtransactions.

6.4 Comparisons to Related Media

Media	Mobile	GPS?	Metadata as Resource	Temporal / Locative Experience	Encoded Data?	Camera?
Barcode Battler	Yes	No	Yes	Somewhat (experience)	No	Somewhat (Barcode)
Geocaching	Yes	Yes	No	Yes	No	Yes
Google Earth	Yes	Yes	Yes	Yes	Somewhat	Yes
I Know Where Your Cat Lives	Yes	Somewhat	Yes	Yes	Yes	Somewhat
Monaco	No	No	Somewhat (Stegano)	Somewhat (gameplay)	Yes	No
Oracle of Seasons / Ages	Yes	No	No	Somewhat (gameplay)	No	No
Pockets Full of Memories	No	No	Yes (Manual Input)	Somewhat (gameplay)	Somewhat	Yes
Pokémon GO	Yes	Yes	No	Yes	No	Somewhat (AR)
Resources	Yes	Yes	No	Yes	No	No
Skannerz	Yes	No	Yes	Somewhat (experience)	No	Somewhat (Barcode)
Spore	No	No	Somewhat (Stegano)	No	Yes	No
The Human Sensor	Yes	No	Somewhat (1 datapoint)	Somewhat (experience)	No	No
WearComp	Yes	Sometimes	Yes	Somewhat (experience)	No	Yes

Figure 37: Properties of select Ludography items related to prototyping objectives

The demonstrative prototype exhibits a combination of aspects in many related media that I have yet to see in concert together. The table in *figure 37* compares what I consider to be defining and effectively-executed characteristics of the final prototype with related works ranging from mobile apps and electronic toys to physical activities and wearables. These characteristics include mobility, locative and temporal elements, GPS and metadata, encoding, and camera functionality, all of which are demonstrated in the multimodal camera and renderer that I have developed. Additional information for all media listed in the table can be found listed in the Ludography in Section 8.2, with web links to informative content. I also discuss additional connections between the prototype and related media in the following sections:

- Section 2.2.1 (Location Data): *Geocaching* and *Google Earth*
- Section 2.3.1 (Metadata in Digital Graphics): *I Know Where Your Cat Lives*
- Section 4.2.1 (Steganography): *Spore* and *Monaco*
- Section 6.1.1 (PCG): *Spelunky*

- Section 6.4.1 (*here*): *Scannerz, Barcode Battlers,* and *Oracle of Seasons*

6.4.1 LCD Barcode Games and Seasonal Gameplay

Figure 38: Barcode Battler II (Epoch Inc., 1991)[54]

Barcode Battlers and Skannerz are LCD games from the nineties incorporating barcode scanners as sources of entropic data, much like the pseudo-random number generators used in PCG algorithms. There is an excellent and inspiring quote by Claus Pias regarding an experience with Barcode Battlers that speaks of the fascination I can envision produced by contextually-relevant gameplay: “how could we possibly know whether cornflakes are more powerful magicians than sweaters or socks? Why should we assume that red wine possesses greater curative powers than a book of poetry by Trakl or a packet of

[54] This image is Copyleft alboran70@yahoo.es. Source: http://20thcenturyvideogames.com/

sparkplugs?" (2007, 231). This ability to maneuver about the physical world and collect data-objects with virtually generated manifestations is highly relatable to the envisioned gameplay mode of the prototype and the curiosity of testing out different datasets for emergent qualities.

Figure 39: Four screenshots of *The Legend of Zelda: Oracle of Seasons* demonstrating season variations (Flagship, Capcom, and Nintendo, 2001)[55]

The seasonal mechanics used in games such as *Banjo & Kazooie*'s Click Clock Wood level (1998) and *The Legend of Zelda: Oracle of Seasons* (2001) allow for interesting mechanics where areas of a level are seasonally blocked off until

55 Image sourced from Wikipedia, contributed by users 'A Link To The Past' and 'David Fuchs'. See https://en.wikipedia.org/wiki/File:Oracle-of-seasons-comparison.png for image and fair-use rationale.

weather conditions make the space traversable or otherwise introduce new challenges to a previously explored location. I have considered incorporating date, temperature and humidity to determine how level hazards are generated for an upcoming prototype build that takes inspiration from these mechanics; the real-world connections make such mechanics ripe for exploration.

6.5 Prototype Outcomes

Developing the prototypes affirmed in my mind that there is value in multimodal images as a resource for application development. They can provide numerous creative design resources in a format that can be easily exchanged between users and non-users alike. Photographers can benefit from the camera prototype's data enrichment and End-Users can enjoy a degree of personalized experience fuelled by their personal curation of generative content. Naturally, the continued usage of these prototypes over time would contribute to a digital archaeology for future extraction by Archivists. Surveillance user groups could very well use such technology as a front for encouraging users to provide data that they could then utilize for profiling or marketing purposes; hopefully, existing digital privacy policies continue to develop and provide agency to consumers to counteract this.

For Application Developers, the outcomes of the prototype represent a vast potential for future creative development with contextual awareness in consideration. While multimodal imagery may not be directly accessible as a ubiquitous technology, the capacity for the datasets utilized in the prototype to be used for developing engaging applications that bridge reality and virtual mobile interfaces is clear from both existing examples such as *Google Earth* and *Pokémon Go* and the work produced in my prototype development. But this does not undermine the value of embedding data into these images in this fashion: the linking of image to contextual data allows for developers to instantaneously access datasets associated with moments that otherwise would be barriered by the development time and know-how required to access, collect and compile the information (as evidenced by my independent development cycle). This contextual data is now in a tokenized format that can be easily exchanged through social networks, and the scripts used to encode and decode the information could likely be developed into an easily-integratable format for future multimodal projects.

To summarize and further evaluate the prototype, I have created the following table based on the key considerations listed in Section 4.1.3 (including Zeng and Qin's principles):

Considerations	**Evaluation**
JSON Format?	Somewhat; I used a Dictionary-object style format, but it could be upgraded to a JSON format or better.
Versioning?	Somewhat; versions are checked and indicated, but there is no method to 'upgrade' old versions of the format yet.
Modular / Genre?	Not yet. I am not sure how best to implement this currently.
Storage & Load Time?	Non-issue. There is not enough data in the current dataset to remotely warrant storage concerns, nor are there any major issues of delays apart from most datasets being reliant on the successful retrieval of a GPS position.
Primary Source Data?	GPS, Date/Time, Camera Image. Secondary Data includes Weather, Celestial information, ArcGIS tables.
Reduced API Dependency?	Yes. The renderer only needed to access the multimodal photos, rather than any data provided by an internet connection or any of the original datasets listed above...
Metadata Element Sizes?	Good. The selected data points had clearly defined limits (e.g. max and min values) and categories, as well as null values for unknown/unrepresentable data.
Refinement / Standardized Data Sources?	Yes. The biome and anthrome data had consistent sources, as did the weather data. The celestial data is pulled from actual astrological calculations. Only the camera image remains unstandardized.
Interoperability / Single, Transferrable Data Unit?	Yes. The multimodal frames are saved as .png images and can be safely transmitted on Facebook and Twitter due to their policies on managing .png images and the steganography encoding technique.
Multilingual?	Somewhat. String data is formatted in English, though the data in the multimodal frames may still be compatible with multilingual systems since they require little to no translation to be used.
Extensible?	No; there are currently no tools to port the information contained in multimodal frames to other metadata standards because the tools have not been programmed and the general lack of ECM in existing metadata standards.

I am excited by the prospect of further development of this prototype and what further insights and creative opportunities I may encounter through its development, especially from a game design perspective. I also wonder about the potential to manipulate the data as a means of 'time travel', to explore how future climate conditions will 'age' the experience or adjustments to expected values might change an experience dramatically, especially as the prototype develops further.

7.0 Conclusions

Please refer to Section 6.5 for a review of the outcomes of the Prototype.

From my understanding of the research that I have conducted, there is a niche in metadata standards for external context that remains to be largely unexplored for digital media files. By integrating such descriptive information into our existing digital images, they become what I have dubbed 'multimodal frames', digital representations consisting of captured and annotative information anchored to the context of the capture itself. The integration of multimodal frames into present day society presents a series of benefits and concerns ranging topics of privacy, representation, gesture, creativity, history, and surveillance that have been explored briefly in the Use Case Discussion chapter. As demonstrated by my prototyping process, the integration of multimodal frames into app development also presents several design considerations.

As I have stated throughout this thesis, my general objective was to provoke a discussion on the subject of multimodalities and external contextualities in metadata associated with media, rather than to put forward definitive rulings regarding how multimodal media should and will be accepted into existing technology, if that does indeed become the case. To accomplish this goal, I developed a multimodal camera capable of creating digital photos embedded with internal and external contextual information and a renderer to exemplify how such data could be creatively presented. I documented my development observations through the process of integrating contextual data elements to my designs and discussed how various user groups could benefit or face challenges from the ubiquitous integration of multimodal frames into modern society. Based on the outcomes of the investigations, the potential usefulness of multimodal photography seems dependent on the level of support it receives for its integration. However, ECM as a concept remains worthy of further discussion as it seemingly represents an area of data representation that has up to this point been largely overlooked, especially for consumer technologies.

Multimodal frames as a concept and the idea of external contextual metadata (ECM) as a new form of metadata for media present a large field of possible departure points for further investigations and examinations. The proposal of integrating ECM into digital objects raises questions for various involved fields including data science, computer science, digital photography, digital anthropology, privacy, cryptography, developers, surveillance corporations, government agencies, archivists, game designers, and others… each with their own respective avenues to pursue. My hope is that this thesis can continue a dialogue towards expanding our expectations for what external contextual metadata can represent and how it can be used in digital applications.

7.1 Further Directions

As a Master of Design student, there were obvious limitations to how much I could explore and prove within this thesis, but the process of investigation revealed further avenues for potential research on the topic of ECM, the relationship between metadata and digital society, and the creative integration of moment-data with rendering software. These opportunities include topics brought up in Section 4.1.3 and the Use Case Discussion in Chapter 5, such as digital histories, creative opportunities, media transferability, surveillance technology and new metadata standards.

While by no means exhaustive, here is a list of additional possible directions that the project and the research findings could take:

- *Multimodal Camera Developer Tools:* I am interested in developing the existing prototype to a level of accessibility that could provide prospective developers the tools they would need to create their own digital applications utilizing ECM or investigate the capture technology itself.

- *Assistive Applications for Photographers and Users:* ECM could be incorporated into existing technologies to provider richer context for photography or photo-viewing applications, particularly for indexical purposes
- *Potential for ECM and Locative parameters in Game Design:* few mobile applications take advantage of hybridized spaces in mobile interfaces (Section 2.4) where real and virtual realities are interlinked; this has been relatively unexplored in Game Design applications and the development of the project toolset could help encourage further explorations.
- *Statistical comparisons with digital photography practices:* given that the practices of digital photography and multimodal photography would theoretically have a large overlap, investigating the statistics behind digital photography would be informative in showing how multimodal photography might be conducted in the future
- *Developing a proper standard for ECM in metadata:* for it to be properly recognized as a category of metadata, ECM would likely need to be brought towards a forum of experts for discussion and analysis concerning its use as metadata, as well as any privacy concerns that might arise with the data integration

- *Non-photographic applications:* this thesis has largely concerned itself with photographic applications of ECM; investigation into other possible capture and use practices would be insightful
- *Improve the fidelity of existing API resources:* in several instances of the prototype's development there were barriers in how accurately data could be collected and represented due to the resolution and availability of relevant data sets.

8.0 Works Cited

8.1 Listing of Works

Adobe Systems, Inc. n.d. "Extensible Metadata Platform (XMP)." Adobe. Accessed November 23, 2019. https://www.adobe.com/products/xmp.html.

Arthur, Charles. 2011. "iPhone Keeps Record of Everywhere You Go." The Guardian. April 20, 2011. http://www.theguardian.com/technology/2011/apr/20/iphone-tracking-prompts-privacy-fears.

Arvidsson, Adam, and Alessandro Delfanti. 2019. *Introduction to Digital Media*. Hoboken, NJ: John Wiley & Sons.

Atkin, Albert. 2013. "Peirce's Theory of Signs." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Summer 2013. Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/sum2013/entries/peirce-semiotics/.

Auger, James. 2013. "Speculative Design: Crafting the Speculation." *Digital Creativity* 24 (1): 1–25. https://www.researchgate.net/publication/263596818_Speculative_Design_Crafting_the_Speculation.

———. 2016. "Speculative Design Is a Counter to Normative Design." Speculative. 2016. http://speculative.hr/en/james-auger/.

"Automaton." n.d. The Merriam-Webster.com Dictionary. Accessed January 3, 2020. https://www.merriam-webster.com/dictionary/automaton.

Backus, Kenny. 2017. "Managing Output: Boredom Versus Chaos." In *Procedural Generation in Game Design*, edited by Tanya X. Short and Tarn Adams, 13–21. Boca Raton, FL: Taylor & Francis, CRC Press.

Bardzell, Jeffrey, and Shaowen Bardzell. 2013. "What Is Critical about Critical Design?" In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 3297–3306. ACM. https://dl.acm.org/citation.cfm?id=2466451.

Barrett, Terry. 1999. "Photographs and Contexts." In *Criticizing Photographs: An Introduction to Understanding Images*, 3rd ed., 96–115. Toronto, ON: McGraw-Hill Humanities, Social Sciences & World Languages.

Batchelder, Ned. 2008. "Spore Creature Creator and Steganography." Ned Batchelder. 2008. https://nedbatchelder.com/blog/200806/spore_creature_creator_and_steganography.html.

Batchen, Geoffrey. 2001. *Each Wild Idea: Writing, Photography, History*. Cambridge, MA: The MIT Press.

Bhanu, K. N., T. Bhaskara Reddy, and M. Hanumanthappa. 2019. "Multi-Agent Based Context Aware Information Gathering for Agriculture Using Wireless Multimedia Sensor Networks." *Egyptian Informatics Journal* 20 (1): 33–44. https://www.sciencedirect.com/science/article/pii/S1110866517301858.

boyd, danah. 2011. "Debating Privacy in a Networked World for the WSJ." danah boyd | Apophenia. November 20, 2011.

http://www.zephoria.org/thoughts/archives/2011/11/20/debating-privacy-in-a-networked-world-for-the-wsj.html.
Braun, David. (2001) 2015. "Indexicals." Stanford Encyclopedia of Philosophy. January 16, 2015. https://plato.stanford.edu/entries/indexicals/.
Bryant, Levi R. 2011. "Introduction: Towards a Finally Subjectless Object." In *The Democracy of Objects*. Open Humanities Press. https://cora.ucc.ie/handle/10468/5639.
Bucklew, Brian. 2017. "Algorithms and Approaches." In *Procedural Generation in Game Design*, edited by Tanya X. Short and Tarn Adams, 271–99. Boca Raton, FL: Taylor & Francis, CRC Press.
Cairo, Alberto. 2016. *The Truthful Art: Data, Charts, and Maps for Communication*. San Francisco, CA: Peachpit, Pearson.
Card, Mackinlay. 1999. *Readings in Information Visualization: Using Vision to Think*. Morgan Kaufmann.
Castaño, Arnaldo Pérez. 2018. *Practical Artificial Intelligence: Machine Learning, Bots, and Agent Solutions Using C#*. Berkeley, CA: Apress.
Chapman, Owen, and Kim Sawchuk. 2015. "Creation-as-Research: Critical Making in Complex Environments." *RACAR : Revue D'art Canadienne / Canadian Art Review* 40 (1): 49–52. https://doi.org/10.7202/1032753ar.
Cohen, S. Marc. 2000. "Aristotle's Metaphysics." Stanford Encyclopedia of Philosophy. October 8, 2000. https://plato.stanford.edu/entries/aristotle-metaphysics/#SubjMattArisMeta.
Crooks, Roderic. 2013. "Review: Mobile Interface Theory: Embodied Space and Locative Media by Jason Farman." *InterActions: UCLA Journal of Education and Information Studies* 9 (1). https://escholarship.org/content/qt20s24769/qt20s24769.pdf?t=mhzvhb.
Deleuze, Gilles, and Claire Parnet. (1977) 2007. "The Actual and the Virtual." In *Dialogues II*, translated by Eliot Ross Albert, 148–52. New York, NY: Columbia University Press.
Desjardins, Jeff. 2017. "What Happens in an Internet Minute in 2017?" World Economic Forum. August 31, 2017. https://www.weforum.org/agenda/2017/08/what-happens-in-an-internet-minute-in-2017/.
Dey, Anind K., and Gregory D. Abowd. 2000. "Towards a Better Understanding of Context and Context-Awareness." In *The PrCHI 2000 Workshop on the What, Who, Where, When and How of Context-Awareness*, 1–12. https://www.researchgate.net/publication/274074382_Towards_a_Better_Understanding_of_Context_and_Context-Awareness.
Dunne, Anthony, and Fiona Raby. 2013. *Speculative Everything: Design, Fiction, and Social Dreaming*. Cambridge, MA: The MIT Press.
Ellis, Erle C., Kees Klein Goldewijk, Stefan Siebert, Deborah Lightman, and Navin Ramankutty. 2010. "Anthropogenic Transformation of the Biomes, 1700 to 2000." *Global Ecology and Biogeography: A Journal of Macroecology* 19 (5): 589–606. https://doi.org/10.1111/j.1466-8238.2010.00540.x.
Evans, Bryce. 2018. "The One Project: Therapeutic Photography with Adults." In *The Handbook of Art Therapy and Digital Technology*, edited by Cathy A. Malchiodi,

282–90. Philadelphia, PA: Jessica Kingsley Publishers.
Facebook, Inc. 2018. “Data Policy.” Facebook. April 19, 2018. https://www.facebook.com/privacy/explanation.
Farman, Jason. 2009. “Locative Life: Geocaching, Mobile Gaming, and Embodiment.” In . Berkeley, CA: University of California. http://www.jasonfarman.com/JasonFarman-Locative_Life_DAC09.pdf.
———. 2010. “Mapping the Digital Empire: Google Earth and the Process of Postmodern Cartography.” *New Media & Society* 12 (6): 869–88. http://www.jasonfarman.com/JasonFarman_Mapping_the_Digital_Empire.pdf.
———. 2014. “The Materiality of Locative Media: On the Invisible Infrastructure of Mobile Networks.” In *Theories of the Mobile Internet: Materialities and Imaginaries*, edited by Andrew Herman, Jan Hadlaw, and Thom Swiss, 45–59. New York: Taylor & Francis Group, Routledge.
Flusser, Vilém. (1991) 2014. “The Gesture of Photographing.” In *Gestures*, translated by Nancy Ann Roth, 72–85. Minneapolis, MN: University of Minnesota Press.
Frith, Jordan. 2015. *Smartphones as Locative Media*. Malden, MA: Polity Press.
Grey, Darren. 2017. “When and Why to Use Procedural Generation.” In *Procedural Generation in Game Design*, edited by Tanya X. Short and Tarn Adams, 3–12. Boca Raton, FL: Taylor & Francis, CRC Press.
Grosz, Elizabeth. 2001. “Lived Spatiality (The Spaces of Corporeal Desire).” In *Architecture from the Outside: Essays on Virtual and Real Space*, 31–47. Cambridge, MA: The MIT Press.
Hammond, Joe, and Paul White. 2015. “Fundamentals of Signal Processing.” In *Fundamentals of Sound and Vibration, Second Edition*, edited by Frank Fahy and David Thompson, 145–211. Boca Raton, FL: Taylor & Francis, CRC Press.
Harper, Douglas. n.d. “Data (n.).” Online Etymology Dictionary. Accessed December 21, 2019. https://www.etymonline.com/word/data.
Herrman, John. 2018. “It’s Almost 2019. Do You Know Where Your Photos Are?” The New York Times. November 29, 2018. https://www.nytimes.com/2018/11/29/style/digital-photo-storage-purge.html.
Holst, Arne. 2018. “Smartphone Industry: Statistics & Facts.” Statista. August 21, 2018. https://www.statista.com/topics/840/smartphones/.
HostingFacts Team. 2018. “Internet Statistics & Facts (Including Mobile) for 2019.” HostingFacts.com. December 17, 2018. https://hostingfacts.com/internet-facts-stats/.
International Press Telecommunications Council. 2015. “Social Media Sites Photo Metadata Test Results.” Embedded Metadata Manifesto. 2015. http://www.embeddedmetadata.org/social-media-test-results.php.
———. n.d. “What Is Photo Metadata?” International Press Telecommunications Council. Accessed November 23, 2019. https://iptc.org/standards/photo-metadata/photo-metadata/.
Iqbal, Mansoor. 2018. “Twitter Revenue and Usage Statistics (2019).” Business of Apps. October 5, 2018. https://www.businessofapps.com/data/twitter-statistics/.
Kopelman, Josh. 2008. “Converting Data Exhaust Into Data Value.” Redeye VC. May 2008. https://redeye.firstround.com/2008/05/converting-data.html.

Larson, Elizabeth, and Richard Larson. 2004. “Use Cases: What Every Project Manager Should Know.” Newtown Square, PA: Project Management Institute. https://www.pmi.org/learning/library/use-cases-project-manager-know-8262.
Letham, Lawrence. 2008. *GPS Made Easy: Using Global Positioning Systems in the Outdoors*. 5th ed. Seattle, WA: Mountaineers Books.
Lin, Ying. 2019. “10 Twitter Statistics Every Marketer Should Know in 2020.” Oberlo. Oberlo Dropshipping. November 30, 2019. https://www.oberlo.ca/blog/twitter-statistics.
Mäkelä, Tapio. 2011. “Memory, History and the Record in the Digital Era: Understanding the Cultural Value of Memory.” In *Euphoria & Dystopia: The Banff New Media Institute Dialogues*, edited by Sarah Cook and Sara Diamond, 140–46. Banff, BC: Riverside Architectural Press.
Manovich, Lev. 2013. “Understanding Metamedia.” In *Software Takes Command: Extending the Language of New Media*, 107–57. New York, NY: Bloomsbury.
Meirelles, Isabel. 2013. *Design for Information: An Introduction to the Histories, Theories, and Best Practices Behind Effective Information Visualizations*. Beverly, MA: Rockport Publishers.
Menkman, Rosa. 2011. *Network Notebook #04: The Glitch Moment(um)*. Amsterdam: Institute of Network Cultures. https://networkcultures.org/_uploads/NN%234_RosaMenkman.pdf.
Mundy, Owen, Shana Berger, Nicole Kurish, Tim Schwartz, and Alissa McShane. n.d. “I Know Where Your Cat Lives: About.” Accessed November 24, 2019. https://iknowwhereyourcatlives.com/about/.
Nourbakhsh, Illah Reza. 2013. *Robot Futures*. Cambridge, MA: The MIT Press.
Noyes, Dan. 2019. “The Top 20 Valuable Facebook Statistics.” Zephoria Inc. November 12, 2019. https://zephoria.com/top-15-valuable-facebook-statistics/.
Oxford. n.d. “Data (noun).” Oxford Learner’s Dictionaries. Accessed December 21, 2019. https://www.oxfordlearnersdictionaries.com/definition/american_english/data.
Perret, Eric. 2016. “Here’s How Many Digital Photos Will Be Taken in 2017.” Mylio: Life In Focus. December 2, 2016. https://focus.mylio.com/tech-today/heres-how-many-digital-photos-will-be-taken-in-2017-repost-oct.
Pias, Claus. 2007. “Play as Creative Misuse: Barcode Battler and the Charm of the Real.” In *Space Time Play: Computer Games, Architecture and Urbanism: The Next Level*, edited by Friedrich von Borries, Steffen P. Walz, and Matthias Böttger, 230–32. Boston, MA: Springer Science + Business Media, Birkhäuser Verlag AG.
Pomerantz, Jeffrey. 2015. *Metadata*. Cambridge, MA: The MIT Press.
Raley, Rita. 2013. “Dataveillence and Countervailance.” In *“Raw Data” Is an Oxymoron*, edited by Lisa Gitelman, 121–46. Cambridge, MA: The MIT Press.
Ramankutty, Navin, and Jonathan A. Foley. 1999. “Estimating Historical Changes in Global Land Cover: Croplands from 1700 to 1992.” *Global Biogeochemical Cycles* 13 (4): 997–1027. https://doi.org/10.1029/1999GB900046.
Rohlinger, Deana A. 2019. *New Media and Society*. New York: New York University Press.
Rosen, Philip. 2001. “Old and New: Image, Indexicality and Historicity in the Digital Utopia.” In *Change Mummified: Cinema, Historicity, Theory*, 225–64. Minneapolis, MN: University of Minnesota Press.

Rubinstein, Daniel, and Katrina Sluis. 2008. “A Life More Photographic: Mapping the Networked Image.” *Photographies* 1 (1): 9–28. https://www.tandfonline.com/doi/full/10.1080/17540760701785842.

———. 2013. “The Digital Image in Photographic Culture.” In *The Photographic Image in Digital Culture*, edited by Martin Lister, 22–40. London: Routledge. https://doi.org/10.4324/9780203797563.

Russell, Archie. 2017. “A Year Without a Byte.” Code.flickr.com. January 5, 2017. https://code.flickr.net/2017/01/05/a-year-without-a-byte/.

Schumacher, Patrik. 2015. “In Defense of Parametricism.” Patrik Schumacher. 2015. https://www.patrikschumacher.com/Texts/In Defense of Parametricism.html.

Shepard, Mark. 2009. “Sentient City Survival Kit: Archaeology of the Near Future.” In *The Digital Arts and Culture Conference, 2009*, 6. Irvine, CA: University of California, Irvine. https://escholarship.org/content/qt4zp0c4x2/qt4zp0c4x2.pdf?t=kx0taz.

Sontag, Susan. (1977) 1990. *On Photography*. New York: Anchor Books.

Stappers, Pieter, and Eliza Giaccardi. 2014. “Research through Design.” In *The Encyclopedia of Human-Computer Interaction*, 2nd ed. The Interaction Design Foundation. https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/research-through-design.

Statista. 2019. “Number of Monthly Active Instagram Users from January 2013 to June 2018 (in Millions).” Statista. December 3, 2019. https://www.statista.com/statistics/253577/number-of-monthly-active-instagram-users/.

Tharp, Bruce M., and Stephanie M. Tharp. 2019. *Discursive Design: Critical, Speculative, and Alternative Things*. Cambridge, MA: The MIT Press.

Thompson, Kristen, Kristen Purcell, and Lee Raine. 2013. “Arts Organizations and Digital Technologies.” Pew Research Center: Internet, Science & Tech. January 4, 2013. https://www.pewresearch.org/internet/2013/01/04/arts-organizations-and-digital-technologies/.

Ware, Colin. 2013. “Foundations for an Applied Science of Data Visualization.” In *Information Visualization: Perception for Design*, 3rd ed. Waltham, MA: Morgan Kaufmann, Elsevier.

Watkins, Ryan. 2016. *Procedural Content Generation for Unity Game Development*. Birmingham, U.K.: Packt Publishing Ltd.

Weibel, Peter. 2009. “The World as Interface - Toward the Construction of Context-Controlled Event-Worlds.” In *Art and Electronic Media*, edited by Edward A. Shanken, 224–27. London: Phaidon Press.

Yau, Nathan. 2013. “Seeing the World in Data.” In *Visual Complexity: Mapping Patterns of Information*, edited by Manuel Lima, 246–48. New York, NY: Princeton Architectural Press.

Yu, Derek. n.d. “What Is Spelunky?” Spelunky World. Accessed February 5, 2020. https://www.spelunkyworld.com/whatis.html.

Zeng, Marcia Lei, and Jian Qin. (2000) 2016. *Metadata*. Chicago, IL: Neal-Schuman.

8.2 Ludography

Davies, Charlotte et al. *Ephémère.* Virtual reality art (1998). (http://www.immersence.com/ephemere/)

Epoch Co. *Barcode Battler*. LCD video game (1991).

Flagship and Capcom. *The Legend of Zelda: Oracle of Seasons & Oracle of Ages*. Video games (Nintendo, 2001). Game cartridges. Nintendo Game Boy Color.

Google LLC. *Google Earth*. Virtual globe (Keyhole, Inc., June 11th, 2001; Google LLC, October 2004). Digital download. Microsoft Windows. (https://www.google.com/earth/)

HHMI BioInteractive. *BiomeViewer*. Virtual globe (HHMI BioInteractive, 2018). Browser application. Microsoft Windows. (https://www.biointeractive.org/classroom-resources/biomeviewer)

Legrady, George. *Pockets Full of Memories.* Installation (2001). (http://www.fondation-langlois.org/html/e/page.php?NumPage=329)

Live Action Role-Play (LARP). Activity (1970s). (https://www.vice.com/en_ca/article/dp545w/a-brief-history-of-larping-1007)

Maksym, Benjamin 'Linker'. *Polymute.* Video game (February 23, 2020). Digital Download. Microsoft Windows. (https://linker.itch.io/polymute)

Mann, Steve. *WearComp.* Electronic devices (1970s-present). (http://wearcam.org/)

Marker, Chris (director). *La jetée.* Film (Argos Films, 1962).

Maxis Games. *Spore.* Video game (Electronic Arts, May 17th, 2008). Digital download. Microsoft Windows. (https://www.spore.com/)

Molga, Kasia. *The Human Sensor.* Wearable technology / artwork (July 23rd, 2016). (http://www.keytoalef.com/kasianet/index.php/the-human-sensor-2/)

Mossmouth. *Spelunky HD.* Video game (October 14th, 2014). Digital download. Microsoft Windows. (https://spelunkyworld.com/)

Mundy, Owen et al. *I Know Where Your Cat Lives*. Interactive map (2014). (https://iknowwhereyourcatlives.com/

Niantic, Inc. *Pokémon GO.* Locative video game (Niantic, July 6th, 2016)

Ollix. *Fog of World.* Locative video game (Ollix, 2012). Digital download. (https://fogofworld.com/en/)

Pocketwatch Games. *Monaco: What's Yours is Mine*. Video game (Pocketwatch Games, 2013). Digital download. Microsoft Windows. (http://pocketwatchgames.com/presskit/sheet.php?p=Monaco)

Radica Games. *Skannerz.* LCD video game (Radica Games, 2000). Electronic device.

Rare, Inc. *Banjo & Kazooie.* Video game (Nintendo, 1998). Game cartridge. Nintendo 64.

Ulmer, Dave. *Geocaching.* Locative activity (2000).

UN3X. *Resources.* Locative video game (UN3X, 2015). Digital download. iOS. (https://www.resources-game.ch/en/)

Vlambeer. *Nuclear Throne.* Video game (Vlambeer, 2015). Digital download. Windows 10. (http://nuclearthrone.com/)

8.3 Assets Used

The following assets were used in the development of both prototypes:

Bellamy, T. I. "Advanced Rule Tiles". *Unity Asset Store*. Unity Extension Asset License. Last downloaded April 3rd, 2020, https://assetstore.unity.com/packages/tools/sprite-management/advanced-rule-tiles-118786

Ellis, E. C., K. Klein Goldewijk, S. Siebert, D. Lightman, and N. Ramankutty. "Anthromes v2.0, data for 2000: ARC/INFO ASCII GRID" (Ellis et al. 2010) & "Base Data for Anthromes v2.0 - Potential Natural Vegetation: ARC/INFO ASCII GRID" (Ellis et al. 2010; modified from Ramankutty and Foley 1999). *Laboratory for Anthropogenic Landscape Ecology.* No license indicated, "for Research Use". Last downloaded March 9th, 2020, http://ecotope.org/anthromes/v2/data/

Fenerax Studios. "Joystick Pack". *Unity Asset Store.* Unity Extension Asset License. Last downloaded April 3rd, 2020, https://assetstore.unity.com/packages/tools/input-management/joystick-pack-107631

Haupt, Hendrik. "Enviro - Sky and Weather". *Unity Asset Store.* Unity Extension Asset License. Last downloaded July 18th, 2019, https://assetstore.unity.com/packages/tools/particles-effects/enviro-sky-and-weather-33963

Hövelbrinks, Stephan 'talecrafter' *et al*. "Aseprite Animation Importer for Unity". *GitHub.* Free-for-use License. Last downloaded Feb 14th, 2019, https://github.com/talecrafter/AnimationImporter

Kula, Suleyman Yasir. "Runtime File Browser". *Unity Asset Store.* Unity Extension Asset License. Last downloaded March 12th, 2020, https://assetstore.unity.com/packages/tools/gui/runtime-file-browser-113006

Nost Games. "Color Palette Swapper Shader". *Unity Asset Store.* Unity Single Entity License. Last downloaded February 12th, 2020, https://assetstore.unity.com/packages/vfx/shaders/color-palette-swapper-shader-153978

Poulsen, Henrik. "SimpleJSON". *GitHub.* MIT License. Last downloaded March 9th, 2020, https://github.com/HenrikPoulsen/SimpleJSON/blob/master/LICENSE

Rainbirth SLU. "Multiple Target Tracking & Framing Camera". *Unity Asset Store.* Unity Extension Asset License. Last downloaded May 19th, 2019, https://assetstore.unity.com/packages/tools/camera/multiple-target-tracking-framing-camera-10554

Signature Group. "CoordinateSharp". *CoordinateSharp.* EULA Non-Commercial Scholarly Usage Grant. Last downloaded March 8th, 2020, https://coordinatesharp.com/

Williams, Jacob 'CheapDevotion'. "Steganography Class for Unity". *GitHub*. WTFPL license. Last downloaded March 9th, 2020, https://github.com/CheapDevotion/Steganography

In addition, the following APIs were called:

OpenWeather. "OpenWeather API: Current Weather and Forecasts Collection". *OpenWeather*. Products and services licensed under the Creative Commons Attribution-ShareAlike 4.0 International license (CC BY-SA 4.0); Data and database open and licensed by Open Data Commons Open Database License (ODbL). Last accessed March 13th, 2020. (https://openweathermap.org/)
IpGeolocation.io "Free IP Geolocation API and IP Location Lookup Database". *IPGeolocation.io*. Terms of service available at https://ipgeolocation.io/tos.html . Last accessed April 20th, 2020. (https://ipgeolocation.io/)

Design Software Used: Adobe Photoshop, Aseprite, Unity, Visual Studio

Appendix A: Accompanying Material

The following material is available upon request from the OCAD University Library:

- digital copies of prototypes, for Windows, MacOS and Android devices
- video demonstrations of Gaia Gate
- design process journals *(see Appendix B)*

Anyone requesting the material may view it in the OCAD University Library or pay to have it copied for personal use.

Appendix B: Process Journals

The prototyping journals are included for completeness and for those interested in a more detailed description of my design process. Please refer to Section 6.0 for a focused summarization of relevant findings towards the thesis objectives. Some edits have been made since their original date of writing. All extensions and assets that were used in the final proof-of-concept and demonstrative prototypes are listed in Section 8.3; any asset mentioned in the journal that is not listed in Section 8.3 was removed from the final product.

The first of these journals include my reflections on the development of the proof-of-concept prototype (including some early thesis musings), whilst the second journal contains my reflections on the development of the final prototype. Due to the combined size of these journals, they have been moved to a separate document, available for access through the OCAD University Library as noted in Appendix A.